V. IMPUTING ACTUAL BEHAVIOR FROM CHOICES MADE UNDER HYPOTHETICAL CIRCUMSTANCES

A. THE ISSUES.

In our earlier (Chapter I and II) overview of concerns/criticisms regarding the accuracy, or interpretative meaningfulness, of value measures derived with the CVM, prominent among those were concerns for biases resulting from the hypothetical nature of the CVM's contingent "market" and the CV payment. Thus, the potential for biases was suggested to result from the fact that the market valuation context, as well as the commodity itself in some cases, will generally be unfamiliar to survey participants; related to the 'unfamiliarity' argument, biases are suggested to be exacerbated by the short time allowed for the valuation process in the CVM relative to the 'weeks or months' 1/ spent by individuals in gathering information -- researching their preferences -- for other, real-life analogous situations. Finally, but related to the above, our earlier overview made reference to research results from cognitive psychologists which suggested the use individuals of heuristic devices in forming judgements in uncertain situations. These concerns share a common theme, viz., a focus on the issue as to how individuals form judgments and values under conditions of uncertainty, or on the question: to what extent can actual behavior be imputed from choices made in hypothetical, uncertain, circumstances?

At the outset it must be re-emphasized that cause-effect statements concerning biases attributable to the hypothetical nature of the CVM have been poorly defined in the literature; in the main, they may be regarded as thoughtful, intuitive, a priori arguments or assertions as to why values derived from the CVM might be biased. Thus, a logically consistent method for organizing and discussing 'hypothetical bias' was not received by the authors. Rather, the authors' initial task was that of attempting to sort through the myriad arguments relating to the substance of hypothetical bias, the time-unfamiliarity issue, as they appear in the CVM literature and the psychology literature concerned with decision-making under conditions of uncertainty, for two purposes: first to set these posited sources for bias in the form of testable hypotheses which relate directly to CVM measures; secondly, to bring together existing evidence which might be relevant for assessing these hypotheses.

These efforts resulted in the following organization for discussions of biases related to hypothetical settings and the CVM. In section B we consider the 'incentives for accuracy' form of the hypothetical bias proposition as it (we argue) relates to hypothetical payment. Bias-related propositions concerning time, preference research and 'unfamiliarity' are assessed in section C. Related to section C's topic, propositions concerning inaccuracies attributable to distorted perceptions of commodities 'traded' in the CVM are considered in section D. Section E addresses the proposition that, with hypothetical goods and payments, CVM values may reflect attitudes as opposed to intended behavior. Our discussions conclude with section F wherein, first, the authors suggest rubrics for issues related to arguments concerning the hypothetical nature of the CVM which might lend clarity and precision to further assessments of these issues and, secondly, results and conclusions from sections B-E are summarized.

Before initiating our analysis, the reader must recognize that results from any one study which has inferential relevance for propositions considered in one section (e.g., time/information issues in C) may also be directly relevant for propositions discussed in other sections (e.g., perceptions and framing of information in D). As implied above, all of this is to acknowledge that may, if not most, of the propositions concerning the extent to which actual behavior can be imputed from choices made under hypothetical circumstances are not distinguishable as separate, independent issues. In treating them separately, the authors do not suggest that they should be distinguishable issues. The partitioning of issues into seperate sections is intended to serve, however imperfectly, the expositional goals of precision and clarity.

B. HYPOTHETICAL PAYMENT: AN INCENTIVE FOR ACCURACY?

As noted above I.C., as well as by Randall et al. (1933), the 'hypothetical bias' notion as it appears in the CVM literature is poorly defined. Too often, the issue is simply described contextually as, for example, "... the hypothetical character of the CV market precludes the derivation of values (which reliable reflect preferences)" (Burness et al., 1983, p. 625). In statements of this form the question is begged as to why the hypothetical market might preclude accurate or reliable responses. On the other hand, one sees in Freeman (1979a) as well as in Feenberg and Mills (1980) a proposition for biases attributable to the hypothetical nature of the CVM which is suggestive of testable hypotheses. Thus, Freeman argues that "In the real world, an individual who takes an action inconsistent with his basic preferences, perhaps by mistake, incurs a cost or a loss of utility. In the (CVM) ... there is no cost to being wrong, and therefore, no incentive to undertake the mental effort to be accurate." (Freeman, 1979b, p. 916)

In its most general from the incentives argument may be re-stated as follows. Let V be an individual's stated valuation for a given commodity X; then the hypothesis consonant with the incentives argument is:

$$V(\text{with incentives}) = V(\text{without incentives}) \qquad (1)$$

As will be argued in Chapter VI, there may be many ways for providing incentives for accurate valuations depending on, among other things, one's criteria for accuracy. In the literature, however, one finds concern with this question limited to one, very specific form of (1) in which the lack of actual payment of 'offered' WTP measures explains the lack of incentives. Effectively then, actual payment = incentives, hypothetical payment = no (without) incentives, and (1) can be rewritten as:

$$H_0: \quad V(\text{actual payment}) = V(\text{hypothetical payment}) \qquad (2)$$

We now inquire as to existing evidence relevant for the form of hypothesis (1) given by (2). The literature abounds with evidence that suggests that (2) be rejected: actual vs. hypothetical payment does result in different choices. Bohm's (1972) seminal experimental work with the CVM, wherein willingness-to-pay values for public television were derived from actual and hypothetical payments, produced results contrary to hypothesis (2) -- actual payments were significantly different from hypothetical payments. From this, Bohm concludes that his results are "... compatible with the general view that that, when no payments ... are involved, people respond in an 'irresponsible' fashion ... this result may be seen as still another reason to doubt the usefulness of responses to hypothetical questions..." 2/ Bohms's findings are supported by results from Bishop and Heberlein's (1979) study of willingness-to-pay/accept for early season goose hunting permits. In comparing 'substantial' differences in willingness-to-accept estimates for hunting permits involving actual ($63.00) and hypothetical ($101.00) payments, Bishop and Heberlein conclude "The stimulus of real dollars ... is simply more powerful than hypothetical dollars ... In plain words, 'money talks' and real money 'speaks louder'

than hypothetical money". (pp. 923-29) As is discussed later in Chapter VI, we note here that Bishop and Heberlein's conclusions in this regard are challenged in a recent paper by Carson and Mitchell (1984). Using alternative (vis-a-vis Bishop and Heberlein) assumptions regarding upper limits for integration and for identifying non-participants Carson and Mitchell demonstrate, using Bishop and Herberlein data, the lack of significant difference between hypothetical and 'actual' payments (p.8). Results from two other sets of studies are relevant for hypothesis (2). First, Coursey et al. (1983) conducted experiments wherein hypothetical and actual willingness-to-accept (WTA) and willingness-to-pay (WTP) measures were related to a subject's tasting of a bitter substance: sucrose octa-acetate. They find a significant difference between WTA and WTP measures when hypothetical, as opposed to actual, payment is involved, a finding explained by the authors as resulting "... mainly from lack of a market-like environment" (15). Secondly, results from tests of actual vs. hypothetical payment on decision strategies reported in the psychology literature 3/ consistently conclude that actual payment makes a difference. Typical of these reported results is Slovic's (1969) conclusion: "It is clear that decision strategies ... differed depending on whether the gains and losses ... were real or hypothetical ... results indicated the importance of committing (subjects) to the consequences of their actions ..." (p. 437)

In contrast to the above, the authors find little if any evidence that would support hypothesis (2). While not directly related to this hypothesis, we find one study which suggest V(hypothetical payment) has predictive value for V(actual payment) in Kogan and Wallach's (1964) conclusion: "It is evident, then, that what an individual does in a hypothetical decision context has some predictive value for a gambling type of task in which decisions represent a firm commitment in a subsequent playoff." (p. 39) Other than this, the authors find but two other studies, the results from which might be inferred as weakly supporting hypothesis (2). These are studies wherein values derived from the CVM are compared with corresponding values derived from the hedonic price method (HPM). These two studies, by Brookshire et al. (1982) and Cummings et al. (1983) are described in some detail below in Chapter VI; thus, in what follows we simply point to the potential relevance of results from these studies to the issue at hand. Such potential relevance must be based on two important assumptions. First, one must accept values derived via the HPM as a measure of actual payment for a commodity -- problems in doing so are detailed below. Secondly, one must accept the argument that individual biases and difference, of the type alluded to above, are immaterial for measures drawn from aggregate behavior -- i.e., at higher level of aggregation, individual biases will generally wash out. 4/ In this regard, one must note the challenges to this argument by Kleindorfer and Kunreuther (1983) as well as by others. 5/ Given these assumptions, comparisons of HPM and CVM (involving hypothetical payments) values may be relevant for assessing (2). 6/ Defining Vh and Vc as values derived from the HPM and CVM, respectively, Brookshire et al. (1982) axiomatically develop the hypothesis $Vh > Vc$; statistical analysis of their data result in their failure to reject this hypothesis. Thus, while not a direct proof of (2), their results can be taken as demonstration of an appropriate relationship between V(actual payment) and V(hypothetical payment): as

measured, respectively, by Vh and Vc, when V(actual payment) should be greater than V(hypothetical payment), this relationship is shown to obtain. Cummings et al. (1983) test the hypothesis given in (2), viz., that Vh = Vc; as in Brookshire et al., their analysis results in failure to reject the hypothesis.

Comparisons aside, the quality of empirical measures of value from the HPM per se are far a level where they might be regarded as accurate, in some sense, estimates for market values attributable to public goods. Thus, results from these comparative studies must be viewed as having questionable weight relative to earlier-described studies in terms of an assessment of (2). Ceteris paribus, one would then tentatively conclude that compelling reasons exist for expecting biases in hypothetical valuations of the sort obtained in the CVM, relative to individual values that would obtain under conditions where expressed valuations must, in fact, be paid. The weight and implications of this tentative conclusion are discussed below.

C. HYPOTHETICAL BIASES RELATED TO TIME.

Consider the following statements of concern about the CVM as expressed by, first, Feenberg and Mills (1980) and, secondly, Bishop and Heberlein (1979).

> "Figuring out what an improvement in water quality of a nearby lake would be worth to you is extremely complex. If it were announced that the lake has been partially cleaned up, you might try it a couple of times, compare it with other lakes, ask friends, and read accounts of the results in the press and elsewhere. Gradually, you would decide the most appropriate modification of your recreational behavior." (p. 60)

> "When people buy things in a market, they may go through weeks or months of considering the alternatives. The process will often involve consultations with friends and may also involve professionals such as lawyers or bankers. It may also entail shopping around for the best deal on the product in question. And, for the majority of items in the consumer's budget, there is a whole history of past experience in the market to base the decision on. All this is markedly different than spending an hour or two at most with a mail survey or a personal interviewer attempting to discern how one might behave in a market for a commodity for which one has never actually paid more than a nominal fee." (p. 927)

These intuitive statements of concern as to the hypothetical nature of the CVM are, in their cited form, obviously not in forms immediately amenable to hypothesis testing. One sees in these statements, however, the strands of an argument which may be stated as a testable hypothesis. At the risk of over-interpretation, the above-cited concerns may be compressed into the argument that individuals require time in order to obtain and mentally 'process' relevant information before informed, 'accurate' judgements can be formed; note here that we beg the question as to whether accurate measures can be obtained with hypothetical payment, regardless of time and information used in the preference research process. If $V(t_o)$ is the expressed value for the CVM commodity X during the typical, short-lived interview used in tile CVM, $V(t_1)$ the value expressed at some later, post-initial interview time, the above arguments suggest rejection of the null hypothesis:

$$H_o: \quad V(t_o) = V(t_1) \tag{3}$$

Variations in (3) could involve obtaining a sequence of values over time wherein endogenous (to the CVM) of exogenous information is made available to or obtained by subjects; if $I_1, I_2, \ldots$ represents increasing amounts of information, such variations would alter (3) as:

$$H_o: \quad V(t_o, I_1) = V(t_1, I_2), \tag{3'}$$

$$t_o < t_1, \quad I_1 < I_2.$$

Cursory inspection of (3) and (3') suggests a number of potentially difficult problems in efforts to test them As an example, across individuals, how does one control for differences in exogeneously-obtained information? Given that 'more' information has qualitative as well as quantitative implications, how does one structure the I's? Most importantly, absent is some notion as to a 'true' value (hypothesis 2) and/or any appeal to reasons why V might converge to some number as t and I become increasingly large: i.e., there is no logical, conclusive way to end the experiment. Surely alternative, better ways exist to draw hypotheses that capture the essence of the 'preference research' problems implicit to the earlier-cited concerns. At a minimum, however, (3) and (3') may serve the purpose of providing a focal point for our inquiry as to the existence of evidence that relates, in one way or another, to the preference research issue.

One finds little evidence in the CVM literature that relates directly to (3) or 3'). Research results do exist, however, that have inferential relevance for these hypothesis. Burness et al. (1983) essentially focus on $V(t_o, I_2)$ in (3') and introduce three techniques designed, in their words, to break '... the hypothetical barrier in CV analysis". (p. 681) These techniques are (i) prefacing willingness-to-pay (WTP) questions with questions regarding the individual's current budget expenditures across six broad budget categories -- after offering a CV value, individuals are then asked where (from which budget category) they will obtain money required to 'pay' the offered value; (ii) after (and before) obtaining a WTP for a specific commodity (an EPA regulation on hazardous waste disposal), other public goods are described to the subject after which the subject may revise his/her WTP measure; (iii) use of the Randall 'bidding game' procedure wherein, after elicitation of an initial WTP 'offer', repeated questions of the form 'would you pay $1.00 more' are asked until the subject indicates: no more (a maximum WTP). Burness et al. find no significant effects on WTP measures resulting from the explicit use of a budget constraint (technique i). a finding which is also reported in Schulze et al. (1983a). The introduction of other public goods (OPG) produces mixed results. The introduction of OPG consistently lowered the offered WTP. In some cases, downward revisions are statistically significant, but in other cases they are not. 7/ Even in cases where lack of statistical significance between initial and OPG-revised bids were found, such results were weakened by large standard deviations and consistent observations of absolute differences in bids of 50% or more. (p. 150) Finally, as in Schulze et al. (1983a) and Desvousges et al. (1983) Burness and his co-authors find that technique (iii) -- use of the bidding process -- significantly affects the WTP measure.

Research results typified by those described above are suggested as relevant for assessments of at least two issues. First, they demonstrate that CV measures are not random numbers: They vary systematically with income, substitute/complementary goods and demographic characteristics as a priori axioms would dictate.8/ Secondly, and of central importance for out discussions, the results are offered as evidence that CV values are individual valuations that reflect a process whereby the subject, in offering a value, has clarified his/her objectives 9/ which is to say that the CV value is a preference-researched bid.10/ That

techniques (i)-(iii) demonstrate a preference researched value is argued to follow from the fact that results from (i) suggest that subjects have considered income - CV commodity trade-offs implied by their offered valuation; results from (ii) may imply that offered bids reflect the subjects' consideration of trade-offs between the CV-commodity and other public goods; and results from (iii) demonstrate that one can, in the CVM, induce subjects to clarify their objectives -- research their preferences -- via the repetitive-question, bidding process.

Obviously, these results have limited, but interesting, implications for (3) and (3'). Formally, techniques (i)-(iii) may be seen as affecting the information term, I, in (3'), where 'more' information is provided by the interviewer (technique ii) or by an induced, introspective process in the case of techniques (i) and (iii). Thus, these data may be seen as relevant for a special case of (3') given as follows.

$$V(t_o, I_o) = V(t_o, I_1). \qquad (3'')$$

When I_1 reflects introspective adjustments to the explicit budget constraint (i), reported evidence suggests a failure to reject (3''). When I_1 reflects information derived from (ii) and (iii), however it appears that (3'') is rejected.

Setting aside estimation problems relevant for tests related to (i)-(iii) 11/ two observations can be made as to how this set of research results relate to assessments of time-related dimensions of the hypothetical bias proposition. First, no objective basis exists for concluding that information effects from(i)-(iii) ultimately result in a 'true' or accurate measure of value. Secondly, the most that one could attribute to the above-cited results is that at t_o (during the interview), values offered by subjects reflect thoughtful consideration of implied trade-offs -- some degree of preference research. But even if this were the case, such evidence would fall well short of speaking to the issue underlying (3') as it is set out by Bishop and Heberlein (1979), Freeman and others, viz., that time per se is required for a meaningfully complete preference research process: values (even with the adjusted information set, I_1) obtained at t_o, $V(t_o, I_1)$, will differ from values obtained at a later period, $V(t_o, I_1)$. This may not always be the case, as is argued by Crocker (1984). In cases where the WTP is an addition to an access fee recently, and actually, paid "much of the environmental and preference information that the respondent had to process in order to arrive at his WTP had therefore already been used by him in his decision to pay the original access fee." (p. 5)

One finds in the literature an abundance of research dealing with learning and 'information processing' capacities of individuals which relates only indirectly to the hypothesis of interest here, but which warrents brief mention. Thus, Kunreuther (1976) and others 12/ suggest that, within the context of high loss-low probability events, serious questions exist as to people's ability to meaningfully absorb -- mentally process -- information. Limited information processing capacity -- causing people to oversimplify problems -- lies at the heart of Simon's (1955) 'bounded rationality' thesis and the 'anchoring' phenomena observed by, among many others, Miller (1956), Ronan (1973) and by Simon and Newell (1971). An understanding of the way in which information is processed by

individuals is seen by Schoemaker (1982) as critical to efforts to predict choice phenomena -- an understanding which is far from complete at the present time.

Brief mention of two additional sets of research results concerning information processing is warranted due to their relevance for future efforts to test hypothesis (3) and (3′). In making decisions under conditions of uncertainty, there exists considerable evidence 13/ that heuristic devices are used by individuals in forming judgements, prominent among which is the 'representativeness heuristic'. This heuristic implies extraordinary reliance on current information irregardless of the quality of such information; prior information is given little weight. With the requisite time differentials in tests of hypotheses related to (3'), the representativeness heuristic suggests the potential for severe problems in controlling/measuring the substance of information changes, I_1 to I_2, and effects of such changes, over the interval t_o to t_1.

Secondly, a number of experimental studies 14/ suggest that, under conditions of uncertainty, individuals may partition, or isolate, decision contexts in curious ways. For example, Tversky and Kahneman (1981) have shown that individuals tend to regard the loss of a $20 theater ticket as more relevant than the loss of $20 in cash, a phenomenon suggesting that individuals mentally partition -- isolate -- groups Of events/actions; i.e., individuals seemingly think in terms of 'mental accounts'. If indeed individuals do consider actions/events/commodities in this isolated, partitioned, mental account context 15/ we know virtually nothing as to how such partitions are formed -- how a mental account is defined. Thus, as examples, one might ask: are mental accounts defined hedonistically (pleasure, pain, aesthetics, etc.), or perhaps functionally (transportation, work, health, etc.)? To the extent that these partitioning contexts are real, potentially serious problems could arise in efforts to test (3') until more is known as to how individuals structure partitions/accounts for obvious reasons: one would be unsure as to the types of information best given to subjects as relevant for approximate real-life information-gathering/processing processes in the t_o-t_1 interval.

From the above we must conclude that little evidence exists that would support or negate hypotheses such as (3) and (3') related to the time-dimensions of the hypothetical bias proposition: the issue remains as an open question. We defer to section F a discussion as to the implications of this void in data for our assessment of the state of the arts for the CVM.

D. PERCEPTIONS, FRAMING AND THE CVM.

There is still another potential dimension of hypothetical bias which relates to the hypothetical commodity 'traded' as a part of the CVM. The relevant line of argument in this regard proceeds as follows. Given that, e.g., environmental changes offered as commodities in many applications of the CVM are hypothetical or, more strongly, imaginary (the subject cannot see or touch the commodity nor, in may cases, can he/she draw on past experience for comparisons of consumption-levels of the commodity), CV measures of value may not be regarded as 'accurate' for two, related reasons: different values offered by different subjects may reflect different perceptions of the hypothetical commodity rather than, as is supposed in the CVM, different preferences; secondly, judgements/values by subjects are dependent on how the commodity is described (how questions are 'framed') and different, in a non-substantive sense, descriptions of the commodity will yield different statements of WTP (value). Concern with this potential source of hypothetical bias is seen, for example, in Schulze, d'Arge and Brookshire's (1981) concern with the need "... to establish a precise contingent market -- the 'good' (commodity) must be well-defined". 16/ Issues related to perceptions and framing are discussed in the following sub-sections.

1. Perceptions. In terms of the 'perceptions' issue one finds in the literature hypotheses concerning how people perceive risky events. It is not clear, however, that the issue is limited in relevance to questions of risk. Consider, as an example, the CV commodity: for a particular river, a change in water quality from boatable to fishable levels. One can only speculate as to the mental image such a hypothetical change might elicit in the mind of any particular subject: an image of 'murky' vs. 'clear' water, or an image of a person sitting in a boat, unused fishing rod in hand vs. the angler fighting a hooked trout on a pristine stream? Surely, this image -- this perception of the CV commodity (or more precisely, of the attributes of the commodity) -- would be relevant for any preference-revealing value offered by a subject. All else equal, the attribution of 'accuracy' to CVM values would then seemingly require a compelling demonstration of at least four relationships: perceptions of hypothetical environmental changes (or changes in availability of any other public good) are in some sense consonant with real effects that would attend the posited environmental change; as something of a corollary to the preceeding issue, subject i's perception of the CV commodity is in some sense consonant with subject j's perception of the commodity -- all subjects are valuing the same commodity; related to the topic of section C, perceived effects (benefits/costs) of the hypothetical commodity are invariant over time (the absence of 'impulse' perceptions); and the independence of perceptions from the quality and quantity of information given to subjects. Thus, as a guide for the discussions that follow, the issues described above are, respectively, described by the following hypotheses.

$$H_o: \quad C(p) = C(a) \tag{4}$$

$$H_o: \quad C(pi) = C(pj) \tag{5}$$

$$H_c: \quad C(t_o) = C(t_1) \qquad (6)$$

$$H_o: \quad C(p/I_1) = C(p/I_2) \qquad (7)$$

where: C = the environmental 'change' used as the CV commodity.
p, a = perceived and actual, substance of the environmental change, respectively.
t_o, t_1 = the time of the CVM interview and some later time, respectively.
I_1, I_2 = distinct information bundles.

Consider first, the hypothesis given in (4) which, essentially, poses the question: are individual perceptions of the substance of a posited environmental change consonant with -- roughly the same as -- the substance of effects that would actually attend the change? As an aside, we note that since such 'substance' is described to individuals as a part of the CVM, in our discussion of (4) the perceptive reader may be troubled by the persistently obvious interdependence between the four hypothesis (4)-(7) and, particularly, between (4) and (7); these interdependencies will be given explicit treatment in later discussions. In term of the limited question posed by (4), however, two sets of issues are of primary interest. The first set concerns the term C(a): the actual substance/effect of a given environmental change. In some cases it may be technically possible to precisely define (estimated) effects that would attend a posited environmental change; as examples: changes in BOD levels in a river; resulting fish populations (by species) and, perhaps, expected catch-rates; changes in TSP or ozone concentrations and changes in visibility. In many other cases, however, the functional relationship between environmental change and the actual effects of such change are not known. 17/ As but a few examples, we know little about household soiling and/or materials damages effects associated with TSP levels 18/; little is understood regarding health effects from air pollution 19/ and we cannot specify risk effects of alternative policies related to the regulation of hazardous waste disposal. 20/In these latter instances, the CVM practitioner has no practical anchor for accuracy. He/she must then rely upon individual perceptions of environmental change-related effects, which then introduces issues related to hypothesis (4), which are discussed below.

In the above described cases where C(a) can be defined, we find in some (but not in others) studies 21/ extensive efforts by the authors to describe the CV commodity (via photographs posters, etc.) in ways (seemingly) designed to bring individual perceptions of the commodity, C(p), in consonance with actual effects that would attend the posited environmental change (our C(a) in (4)). We do not find, however, evidence that the authors attempted to test the effectiveness of their efforts in this regard, i.e., the authors do not address hypotheses of the sort typified by (4). Rather, the consonance of C(a) with C(p) is simply asserted, as in the following (relevant editorial questions in parentheses): "The (water quality) ladder's major attribute is that it easily establishes (in the minds of individuals?) linkages between recreation activities and water qualities ... it directly introduces the relationship between (the individual's perceptions of?) activities and (the individual's perceptions of?) different water quality levels ..." (Desvousges et al., 1983, pp. 4-11); "... bids were solicited for the same well-defined public good,

visability at the Grand Canyon National Park. Specification of this good -- implicitly, C(a) vis-a-vis C(p) -- was assured (emphasis added) by presenting all respondents with the same set of photographs of known visibility levels...". (Schulze et al., 1983a, p.2-2)

In terms of the second major set of issues relevant for assessment of hypothesis (4), assume that C(a) is known and that it can be 'adequately' described. We now inquire as to results from experimental/empirical research which directly relate to (4). We find such evidence only in the literature concerning decision-making under conditions of risk and uncertainty. In this regard, Slovic and Tversky (1974) report results from a study wherein subjects were confronted with various paradoxes; after making their choices -- reflecting C(p) -- they were given an authoritative argument against their choice -- a representation of C(a). Most subjects did not change their particular choices. Implications of findings such as this are summarized by Slovic et al. (1980) as follows: "A great deal of research indicates that, once formed, people's beliefs change very slowly, and are extraordinarily persistent in the face of contrary evidence ... New evidence appears reliable and informative if it is consistent with one's initial belief, whereas contrary evidence is dismissed as unreliable, erroneous or unrepresentative." (p. 189) Thus, given an accurate description of C(a) to individuals interviewed in the CVM, substantial evidence suggests, in terms of risky/uncertain events, the rejection of (4); an effort to adopt economic models to reflect such behavior, described as 'cognitive dissonance', can be seen in the work by Akerlof and Dickens (1982). We do not find such evidence related to non-risky events; to the extent that the risky-event evidence can be generalized, however, rejection of (4) implies that variations across individuals of CVM values may reflect differences in perceptions of the hypothesized commodity. Finally, we note the relevance for the issue as to how individuals perceive C(a), of the literature that suggests that individuals have a 'threshhold' of sensitivity. 22/ Thus, individuals may be insensitive to CVM commodities that represent 'moderate' environmental changes, and react (in a valuation sense) only to changes involving extremes, for example, eutrophication vs. pristine lake conditions. The result of such behavior is often reflected in increasing marginal value functions (Crocker and Forster, 1984). "Threshhold" phenomena are seen, for example, in the works of Crocker, Dauber and Young (1981) as well as in Loehman et al. (1979).

Referring now to hypothesis (5), a recurring theme in the discussions above -- all subjects perceive the same commodity -- was that with or (arguably) without the standard C(a), variations in perceptions across individuals may severely weaken the meaningfulness of CV measures inasmuch as individual values would be attributable to different commodities. In instances where C(a) cannot be estimated, as noted above, the CVM practitioner may be tempted to rely on individual perceptions of the commodity, in which case comparable perceptions of the commodity by all subjects -- hypothesis (5) -- becomes particularly important. We then inquire as to the nature of available evidence related to hypothesis (5).

Indirect evidence related to (5) is found in the above-cited works by Slovic and others. For example, Slovic et al. (1980) find systematic differences in the perceptions of a given activity between groups of laypeople, groups of experts and between experts and laypeople (p.211). We find in one CVM application, however, information which directly relates to (5). Cummings et al. (1981) used the CVM to estimate benefits

attributable to reduced household soiling which was, in turn, attributable to reductions in TSP concentrations. The researchers were unable to specify a relationship between lower TSP concentrations and reductions in household soiling (C(a) was unknown). 23/ Therefore, following a qualitative explanation to subjects of the TSP-household soiling relationship, WTP measures were obtained for alternative percentage reductions in TSP concentrations, leaving to individuals the (perceptive) task of translating reductions in TSP concentrations into reductions in household soiling. Prior to the WTP questions, subjects stated the number of hours/week that they spent in household cleaning activities (W). Following the WTP question, subjects were asked how they expected W to be affected by the posited change in TSP concentrations; i.e., for the posited environmental change to which their WTP applied, they were asked their perception of the work savings (WS) that would attend the environmental change. Implicitly, for each individual i in the Cummings et al. (1981) survey, WS (i) may be viewed as a measure of C(p) in (5). WTP measures were regressed against the WS variable and the WS variable was found to be statistically significant -- WTP measures offered by individuals varied systematically with individual perceptions of WS: individuals had significantly different perceptions ($C(pi) \neq C(pj)$ in (5)) and valued differently perceived WS's differently. Thus, with C(a) known, and particularly with C(a) unknown, available evidence suggests significant differences in individual perceptions of uncertain and, perhaps, unfamiliar commodities.

Hypotheses (6) and (7) involve, in large part, issues discussed above in section C. Therefore, aside from two observations of particular interest to the perception questions at issue here, time-information problems will not be belabored in this section. We should comment, first, on the (perhaps inextricable) interdependencies between (7) and (4) (and, to a lesser extent, (6)) and between (7) and (5). Obviously, the provision and 'processing' of information -- the substance of hypothesis (7) -- is of central importance to empirical tests focused on (4) and/or (5). For example, C(a) is established by giving the subject information. In this regard, questions related to (7) include: what kind and how much information? A second, but related observation concerns the substance of information -- 'substance' as opposed to how questions are asked (framed), an issue to be discussed below. Referring to 'information bias', Randall et al. (1983) consider the argument that '... variations in the materials describing contingent markets may influence (WTP responses)" (p. 641). In this regard, they contend that CVM demonstrations that WTP values vary with information/materials may not be evidence of any kind of bias. Rather, if alternative materials/information given to subjects are relevant to the choice problem, "... information that changes the structure of the market should (arguably) change the circumstantial choices made therein" (p. 641). It is not clear exactly what Randall et al. have in mind in referring to information that 'changes the structure of the market'; but materials/information describing the CV commodity is seemingly included. This statement then invites the following interpretation which is relevant for (3') as well as (7): information that affects -- changes -- an individual's perceptions of the commodity should change the individual's valuation of that commodity. In examining the implications of this interpretation of Randall et al.'s argument, it is understood that this

is not necessarily their interpretation; while several interpretations are possible, the one which best fits the context of their arguments is examined below in our discussions of framing issues. This interpretation, however 'strawman' in nature it might be vis-a-vis Randall et al.'s intended interpretations, is useful, in addressing a potential source for confusion in assessments of hypothetical bias.r

If one ties perception to preferences and tastes, the line of logic: "different information implies different perceptions, preferences and tastes implies different valuations" has clear appeal in its consistency with utility theory. An important distinction arises, however, in using the market analogy to argue that this logic suggests 'no bias' in CVM measures. In the market, at any instant in time market valuations cut across, in some average sense, individuals with heterogeneous information states reflecting, among other things, different experiences/histories with the commodity and differing levels of effort (across differing time-spans) in acquiring/processing information; 'new' information can then be expected to affect valuations much more slowly and, as suggested in the following, to have small relative effects. In the CVM, however, in the many applications wherein individuals are basically unfamiliar with the environmental commodity, particularly as it is viewed in a market context, the initial -- at the interview -- set of information is the same for all individuals and, plausibly, the variance of individual past experience/history is very small relative to market goods. Thus, changes in information, and particularly changes in time available to process information, can be expected to have valuation impacts not at all analogous to the market. In the case of the CVM, market-like heterogeneity in terms of individual preferences, tastes, experiences, etc., as would be reflected in market prices, can be expected only after considerable variation of I in (3') and (7) as well as with variation in t_1 -- time with which to process -- as each individual chooses -- the information.

2. Framing.

The second major set of issues relevant for assessments of potential biases brought about by the fact that the CV commodity is a hypothetical commodity concerns the argument that values may be affected by the way in which the market context and/or WTP questions are framed -- how they are described to the individual. Formally, if D1 and D2 are different, but 'true' or accurate, descriptions of the same commodity and V is the CV value offered for the commodity, then the hypothesis of interest here is given by

$$V(D1) = V(D2) \qquad (8)$$

It is understood, of course, that perceptions affected by D1 and D2 underline the valuations V. In the following descriptions of research results relevant for an assessment of hypothesis (8), we consider this issue as it relates to two, obviously related, settings: first, D1 and D2 reflect alternative decision (market) contexts and, secondly, D1 and D2 are alternative ways of framing the WTP question within the same decision/market context.

3. Framing Decision (market) Contexts. A large number of studies have been conducted concerning the effects of context -- words used in describing decision alternatives -- on choices/decision-making

(Schoemaker, 1982). The focus of a large part of these studies is the extent to which individual behavior under conditions of uncertainty, is consistent with predictions drawn from expected utility theory. In this specific regard (comparisons with expected utility theory) we simply note Arrow's (1982) conclusion concerning the case being made "... for the proposition that an important class of intertemporal markets shows systematic deviations from individual rational behavior ..." (p.8) Our present interests are in results from that part of the 'decision-making under uncertainty' literature that relates directly to hypothesis (8). Two examples can serve to typify the general nature of experimental results relevant for this issue.

First, Tversky and Kahneman (1981) conduct an experiment wherein subjects are asked to consider two programs, programs A and B, which are designed to mitigate the effects of an outbreak of an unusual Asian disease which is expected to kill 600 people. The consequences of adopting A or B are described in two, effect-equivalent ways:

A : exactly 200 people will be saved.
B : 1/3 probability of saving all 600 people,
2/3 probability that none of the 600 are saved.
A' : 400 people will die.
B' : 1/3 probability no one will die,
2/3 probability all 600 people will die.

For (158) subjects given alternatives A,B, 76% chose program A. For similar subjects (169) given alternatives A',B', 87% chose alternative B'. Thus, individual choices between alternatives were, seemingly, substantively affected by framing the same alternatives with the context of lives saved as opposed to the 'dying' context.

Similarly a second study by McNeil et al. (1932) involved comparisons between two therapies for treating certain forms of cancer: surgery and radiotherapy. Different groups of individuals, including a group of physicians, were given one of two sets of information:

(1) probability of survival with surgery (for 1 and 5 years)
(2) probability of survival with radiotherapy (for 1 and 5 years)
(1') probability of dying within 1 and 5 years with surgery
(2') probability of dying within 1 and 5 years with radiotherapy

Probabilities in 1 (2) were one minus the probability in 1' (2'). 86% of the group of physicians given alternatives 1-2 preferred surgery (alternative 1); only 50% of the physicians given alternatives 1'-2' preferred surgery, however. As in our first example, choices are seen to be affected by differences in dying-survival contexts within which alternatives are framed.

Demonstrations of framing effects on individual choices are not limited to stark contexts involving life or death; such effects are demonstrated for choices involved in gambling and in the purchase of insurance against monetary hazards. 24/ We do not, however, find demonstrations of this type of framing phenomena applied to decision settings wherein some sort of risk per se is not the central issue. Thus, the extent to which the

above-reported results imply a general rejection of (8) is simply not clear. We return to this issue at the end of this subsection.

In addition to the above, one finds in the CVM literature results which relate in an interesting way to the framing hypothesis given in (8). In section V.C. and Chapter III's discussion of potential biases related to time and 'preference research' issues, results from one set of CVM experiments were offered as relevant for assessing the extent to which WTP measures derived in the CVM were, in some sense, preference-researched values or, at a minimum, indicative of the non-randomness of CV measures (see, particularly, section V.C above). This experiment set involved comparisons of CV measures when the commodity is valued alone with those obtained when the same commodity is valued within a context where other public goods are discussed. 25/ As discussed earlier (section V.C), results from these experiments were only weakly relevant in speaking to hypothesis (3') wherein time in the preference research process was of central importance. These experiments, as well as their results vis-a-vis the preference research hypothesis (3'), can be seen as relevant to our present discussions inasmuch as they demonstrated that values for a commodity, when the commodity was framed/described in isolation -- D1 in (8) -- differed from values for the same commodity when the commodity was framed/described within a context that included other public and/or private) commodities -- D2 in (8). With this context as a means for testing hypothesis (8), the finding $V(D1) \neq V(D2)$ is reported for an air quality commodity by Schulze et al. (1983) for a 'hazardous waste regulation' commodity by Burness et al. (1983) and for a public facilities (park system) commodity by Majid et al. (1983)

Recall now the earlier-cited assrtion in Randall et al. (1983) (in the balance of this argument, simply 'Randall') that "... information (read: framing) that changes the structure of the market should (arguably) change the circumstantial choices made therein". (p. 641) While 'framing' in the sense of word/probability substitutions (e.g., probability of death vs. probability of survival is not easily viewed as a change in the structure of the market, one might, and Randall seemingly does 26/ view contextual changes of the 'other goods' stripe as effectual changes in the market structure; if this view is defensible, above-described results do not directly imply framing-related biases in reported CVM measures: V(D1) 'should' be different from V(D2). In terms of decision-making under uncertainty, received theory 27/ assumes that all possible choices, states of the world and consequences (vis-a-vis states of the world) of actions are certain and known by individuals. 28/ A simple application of this assumption, an extension of the more general assumption of rationality basic to economic theory, would lead us to reject the above interpretation 29/ of Randall's 'arguable' proposition. Thus, since individuals know -- are perfectly aware of -- the dimensions of all 'other public goods' (the contextual frame D2) then, ceteris paribus, individual choices regarding one specific public good should be unaffected by whether or not (redundant) information regarding other public goods is made available; the reported findings $V(D1) \neq V(D2)$ must then be 'explained' on grounds other than changes in market structure -- framing bias may be one such ground.

However, there are at least two reasons for questioning the position outlined above and, by implication, for imputing some weight to Randall's argument. First, for decisions involving uncertainty -- and decisions elicited

in the CVM surely involve uncertainty -- the nationality assumption in general, and the assumption of certain, comprehensive knowledge of choices, states and consequences in particular, are widely questioned as to their empirical validity (Shoemaker, 1982). Indeed, as discussed above in V.C., the mental capacity of individuals to 'process' but a very limited amount of information is suggested by results from a number of empirical studies. As an example, where C and S refer to choices and states, respectively:

> "As far as C is concerned, it does not require much ingenuity to think of decision problems in which the essence of the problem is that one does not know what options are available. As far as S is concerned, it is easy to think of examples in which one cannot list all possibilities that may occur (And, of course, knowledge of S implies that no one is ever surprised: is this the case in real life?)." 30/

Secondly, appealing to the 'familiarity' arguments discussed above in V.C., and accepting the assumption that individuals are reasonably cognizant of choices in their consumption set, one might argue that the CVM involves, in most applications, what is essentially the introduction of a 'new' commodity to the individual's consumption set. Given that the commodity is hypothetical, and recalling earlier discussions of perceptions, new information/materials may alter the 'shape' perceptions of the new commodity, giving rise to what would indeed be a meaningful 'change' in the commodity (a la Randall, a change in the structure of the market). It must be noted, however, that this argument may suggest, among other things, that the CVM may produce a decision 'climate' rich in its potential for confusion.

To briefly summarize, while a strong case is found for the argument that the framing (wording) of decision contexts can affect individual choices in some settings -- settings wherein some form of risk is of primary importance -- the implications of this argument for hypothesis (8) as it relates to an assessment of the CVM are not clear. For applications of the CVM to environmental commodities, analogies to the 'death-survival' examples are not immediately obvious. Possible analogies might be: increased visibility vs. reduced haze; increased water quality vs. reduced pollution; but these analogies are imperfect at best. While results that might suggest rejection of (8) are weak, research results that might suggest acceptance of (8) are weaker still. Such 'evidence' per se is non-existent. All that we have are arguments with questionable appeal as to why CVM-study results that suggest rejection of (8) might be interpreted differently. Thus, we can say little more than that the case for or against the potential for biases emanating from the framing of market contexts remains as an open empirical question.

b. Framing The WTP Question. In preceding discussions, our focus on market 'structure' or context was, more precisely perhaps, a focus on the framing of the CVM commodity. In the death/survival examples, alternative 'choices' are analogous to the alternative 'commodities' in the CVM. In those experiments, however, there is nothing analogous, in terms of the framing problem 31/ to the hypothetical WTP question posed in the CVM. Thus, while the WTP question -- the CVM's counterpart to a market price -- is obviously a part of market structure per se it is treated separately

here inasmuch as evidence available for assessing the framing bias hypothesis (8) as it applies to the hypothetical WTP question is distinct from that relevant for assessing (8) vis-a-vis the hypothetical commodity.

We have made repeated references to the confusion that one encounters in the CVM literature arising, in large part, from imprecise rubrics for sources of potential biases; see particularly, our earlier (Chapter II) discussion of the many 'faces' of the hypothetical bias proposition. In Chapter III, reference was made to concern in CVM studies with biases emanating from (i) the payment vehicle, (ii) starting points, and (iii) preference research (as addressed via the explicit use of 'budget constraints). 32/ Given that (i)-(iii) directly relate to the question as to how WTP measures are affected by the manner/context in which the WTP question is framed, it may be convenient to view these sources of bias within the rubric of framing bias; convenience aside, results from CVM experiments regarding (i)-(iii) are of obvious relevance for our assessment of (8) as it relates to the WTP question.

Given the extensive discussions of CVM studies and experimental results related to (i)-(iii) in Chapter III, our present purposes are adequately served by a brief review of those results (Schulze, 1981, Rowe and Chestnut, 1983); regarding (iii), we simply note in passing the potential relevance of the 'unfamiliar commodity' and Randall's 'materially-changed market structure' arguments, and the resulting conundrum, for evidence derived from this set of experiments. There have been a number of CVM experiments which focused on issues (i)-(iii). While it is no surprise that unanimity does not exist as to the interpretations of results from these experiments, the following generalizations appear (to the authors) to be reasonable. Referring to (i), tests for 'vehicle bias' have focused on the sensitivity of WTP measures to descriptions (framing) of the method of payment: common examples of payment methods used in these studies are higher tax payments, higher utility bills and higher prices for goods and services purchased. Four out of five studies 33/ found significant effects on WTP measures attributable to the way in which WTP questions were framed vis-a-vis the payment mechanism; obviously, such evidence suggests rejection of (8). Referring to (ii), there appears to be general consensus that WTP questions framed within the context of a 'starting point' -- an initial value; e.g., 'would you be willing to pay $10.00?' -- results in biased measures. Since about 1980, CVM researchers have, therefore, followed the lead of Mitchell and Carson (1981) in using 'payment cards' -- the individual is given a chart on which is written many different values (e.g., from $.50 to $50.00 in increments of $.50) and is asked something like '... referring to this chart, what is the maximum amount that you would be willing to pay ...?". While demonstrative of the fact that the 'starting points' result in framing-type biases, the issue per se may now moot given that 'starting points' are seemingly no longer used in applications of the CVM. Finally, referring to (iii), it would seem that WTP measures are unaffected by whether or not the WTP question is framed within a context where the individual's budget (income, present allocation of income across expenditure categories, and expenditure category(s) to be reduced for 'payment' of the offered WTP) is explicitly considered by the individual in offering his/her WTP. One caveat is relevant in this regard, however: there exists one demonstration that the manner in which budget information is presented (framed) may affect the WTP response

(Schulze et al., 1981).

By way of a summary, there is a good deal of evidence that suggests the potential for biases in CVM measures resulting from the framing -- description -- of commodities and payment mechanisms as well as from distorted perceptions of commodities (as described to individuals). As noted earlier, it may be possible to develop means for including perception issues in economic models from which testable hypotheses are derived; examples in this regard are seen in the works of Akerlof and Dickens (1982) as well as in Coursey et al. (1983). On the other hand, framing issues present a different problem. As noted by Shoemaker (1982), objective assessment of this potential is made difficult by the fact that "... problem representation is inherently a subjective matter, (therefore) it is subject to only limited normative evaluation. Indeed, there exists no general normative theory as to how problems should be defined, or how language and context should be encoded." (p. 556) Notwithstanding the lack of a normative theory to guide assessments of framing-type biases, general guidelines for framing questions do exist, as will be discussed below in section E. We defer to section F a discussion of the implications of these issues for our state of the arts assessment of the CVM.

E. ATTITUDES VS. INTENDED BEHAVIOR.

Given the hypothetical, 'artificial' (Bishop et al., 1983) structure of the CVM, Bishop and Heberlein (1979) have suggested that measures derived by the CVM may reflect individual attitudes vis-a-vis (e.g.) an environmental commodity as opposed to intended behavior (a meaningful intention to actually pay the stated WTP). Their proposition, which draws on works by Schuman and Johnson (1976) focuses attention on questions related to the causal chain -- attitudes-intended behavior. Thus, at issue are the questions: are attitudes indicative (good predictors) of intended behavior; is intended behavior indicative (a good predictor) of actual behavior?

In one's reading of the attitude-intended behavior controversy as it appears in the psychology literature 34/, one might be tempted to argue that the power of responses to attitudinal questions for predicting intended behavior is of no, or questionable, relevance for the CVM inasmuch as questions posed in the CVM are (or should be) well-framed questions about intended behavior per se: questions about attitudes are not asked in the CVM, ergo, attitude-behavior issues are not relevant, Q.E.D. This line of argument is implicit to Randall et al.'s (1983) rejection of the relevance of the attitude-behavior issue. (also see Rowe and Chestnut, 1982). After reviewing the Schuman-Johnson and Ajzen-Fishbein papers, the authors find compelling Randall et al.'s argument as to the questionable relevance of the attitude-behavior issue for the CVM, particularly in light of the comforting assurances by Ajzen and Fishbein that the potential for attitude-related biases can be mitigated by questionnaire designs wherein close consonance is established between actual and hypothetical situations via describing intended behavior in terms of specific actions, contexts, targets and time frames. (Ajzen and Fishbein, 1977, pp. 888-9) Thus, it would seem, the hypothetical question posed to restauranteurs in LaPiere's (1934) seminal work concerning attitudes and behavior "Will you accept members of the Chinese race as guests in your establishment?" elicits an attitude; intended behavior is elicited by posing -- framing -- the question as, e.g., "Will you receive and serve Chinese guests, Messrs. Lin and Chow (here is their photograph), at table number 12 tomorrow afternoon at 1:15 p.m.?".

The notion that attitudinal questions elicit attitudinal responses and questions as to intended behavior elicit behavioral responses, regardless of whether the behavior at issue is hypothetical, may be seen as consistent with results from empirical studies concerning the 'preference reversal' phenomenon. 35/ When asked (relatively) attitudinal questions regarding preference between bets, subjects made choices inconsistent with predictions for expected utility (EU) theory. When then asked what they would pay to participate in a bet, subjects reversed their decision (reversal of preference), and made choices consistent with EU theory; such reversals were found to occur when payment was real or hypothetical (also, see Schoemaker, 1982, pp. 553-554). Thus, behavior-based questions elicited "... the right answer ..." (Randall et al., 1983, p. 638) while attitudinal questions did not. An obvious caveat applies to this conclusion. The standard for a 'right answer' in this context is behavior deduced from EU theory and, as discussed above, the relevance of EU theory in predicting real world decisions is widely challenged.

Thus,in response to Bishop and Heberlein's suggestion that the CVM may elicit attitudinal responses as opposed to willingness-to-pay in the sense of intended behavior, the following observations are relevant. First, purely attitudinal questions may perform poorly as indications of intended behavior. Secondly, some evidence, albeit challengeable evidence, exists which supports the argument that questions about intended behavior may yield accurate predictions of behavior. Third, criteria exist (Ajzen and Fishbein) for mitigating attitudinal biases in responses to questions concerning intended behavior; we note, however, the lack of definitive evidence that adherence to Ajzen and Fishbein's criteria will necessarily eliminate attitudinal biases (we also note the lack of guidelines for judging what 'adherence' might mean).

We wish to close this section by providing some context for the Ajzen and Fishbein (A-F) criteria for mitigating attitudinal biases. This context is provided via an example of a CVM study wherein A-F criteria were applied in the questionnaire design process. Consider the context of the WTP question used in Desvousges, Smith, and McGivney's (1983) (DSM) earlier described study of water quality (also in this regard, see the study by Crocker, 1984). Following A-F's criteria for specific context, targets, actions and time frames, prior to posing WTP questions, DSM ask individuals earlier, specific instances when the individual has visited specific places along the Mohongahela River for recreational purposes: 'your actual use' of recreational areas in the River is established in the individual's mind. The structure of their questions as to intended behavior is as follows: (Appendix D, pp. D-7 to D-13)

specific context "keeping in mind 'your actual use' of recreational areas along the Mohongahela River ..."

specific action/time frame "... what is the most that you would be willing to pay each year (time frame) ..."

specific action "... pay in higher taxes and prices for products that companies sell ..."

specific target "... to raise the water quality level in the Monongahela River from x to y".

In the above, it is interesting to note that the device used by DSM to enhance the specificity of actions -- higher prices and taxes -- introduces the potential for framing biases of the 'payment vehicle' type discussed above in V.D.2, a potential seemingly viewed as a blessing by DSM, e.g., "This payment vehicle was selected because it corresponds with how people actually pay for water quality (do subjects know this?), connotes no implicit starting point, and produces a vehicle that will bias the response downward (emphasis added), if in any direction, because of public attitudes towards increased taxes and higher prices" (p. 4-16). In conclusion, we note in passing that in DSM's comparisons of CVM values with values derived from the TCM (discussed below in Chapter VI) we will see that above-cited anticipation of underestimations in CVM measures attributable to framing biases are apparently forgotten in their value-comparison analysis.

F. CONCLUDING REMARKS.

In this chapter an effort has been made to organize, discuss and assess the many potential sources for bias in CVM measures that derive, in one way or another, from the hypothetical nature of the CVM's commodity, market and 'payment'. In cases where a set of intuitive arguments lend themselves to more precise representation as one or more statements of hypotheses, general hypotheses are offered as a tool for providing focus to an assessment of the arguments. Major sets of biases related to the hypothetical nature of the CVM and, when appropriate, null hypotheses related to them which were developed in this chapter; these null hypotheses are summarized as follows. In what follows, HB, hypothetical bias, is understood to conote the proposition: "Hypothetical bias (in the CV measure) may result from the fact that:".

HB.1 Payment in the CVM is hypothetical.
V(actual payment) = V(hypothetical payment).

HB.2 The CVM Commodity is hypothetical;

HB.2(a) This is to say that preference research for the unfamiliar, hypothetical commodity takes time
$V(t_o) = V(t_1)$ and/or

HB.2(b) This is to say that preference research for the unfamiliar, hypothetical commodity requires information and time to process the information.
$V(t_o, I_1) = V(t_1, I_2)$ and/or

HB.2(c) This is to say that:

(i) individual perceptions of the CV commodity will not be consonant with the 'actual' commodity offered, C(p) = C(a), and/or

(ii) given a description of the hypothetical commodity, different individuals will perceive and, therefore, value, different commodities. C(pi) = C(pj), and/or

(iii) commodity perceptions, and therefore values, will change with the passage of time and/or the accumulation of information.

HB.3 Payment and the Commodity are hypothetical.

HB.3(a) Therefore, WTP measures will be affected by the context within which the commodity and payment is described, or or framed.

V(D1) = V(D2) and/or

HB.3(b) Therefore, the CVM will elicit responses reflecting attitudes rather than intended behavior, and attitudes do not perform well as indicators of intended behavior.

Subsumed in this structure for assessing potential biases in CVM measures attributable to the hypothetical nature of the CVM are sources for bias described in earlier works under the rubrics 'vehicle bias', 'starting point bias', 'information bias' and 'hypothetical bias'.

Based on our assessments and discussions of research results drawn from the literature as they relate to HB.1-HB.3, three general observations seem apparent in terms of implied tentative conclusions regarding the state of the arts of the CVM; common to all three observations must be the understanding that, as reflected in CVM experiments conducted to date, researchers have only recently begun to address several empirical questions that must be viewed as fundamental to any demonstration which purports to establish, in a compelling way, that the CVM can be designed in such a way that meaningful values are derived. First, we observe that the framing questions underlying HB.3 imply the need to rationalize and apply to questionnaire design, criteria (perhaps) of the sort set out by Ajzen and Fishbein for eliciting values which (all else equal) reflect behavioral intentions. Obviously, this will be no mean task; this is particularly true for efforts to rationalize criteria in the sense of establishing standards by which the investigator can empirically test the extent to which the CVM design approximates 'actual conditions'. Other related fundamental questions which remain unanswered by experimental research are those related to time and (perhaps inextricably) information -- HB.2(a), (b), (c.iii). Given, in many applications of the CVM, the lack of congruence between people's experiences and the hypothetical commodity, as well as the hypothetical market context within which the commodity is to be valued by them, one cannot easily dismiss the intuitive appeal of the ('familiarity') argument that information processing, which involves the introspective process of examining -- researching -- one's preferences, will take different forms -- and, therefore, yield different value responses -- over different time frames. While certainly challenging, these framing and time/information issues do not, in the authors' minds, pose impossible question; i.e., implied questions are amenable to statements in the form of testable hypotheses. At this point at least, the relevance of these issues for one's assessment of the CVM is an indication of ignorance -- unanswered questions -- as opposed to a definitive indication of unresolvable weaknesses in the CVM.

Secondly, experimental applications of the CVM to date have yet to address in a compelling way, the question as to the extent to which individual perceptions of the hypothetical commodity -- the item which they are asked to value -- are in any sense consonant with the actual commodity offered in the CVM; in this regard, we note occassional confusion in CVM studies as to the 'commodity' relevant to the valuation decision 36/ and the relevance of framing issues for efforts to empirically address the perceptions issue. At a minimum, this question appears to be amenable to empirical inquiry. Such is not the case in instances where actual effects of (e.g.) an environmental change cannot be specified. In such cases, one

cannot define a standard against which to assess commodity perceptions by individuals. Therefore, we must conclude that use of the CVM for deriving individual values for such commodities will be an empty exercise given that one cannot distinguish between value differences (among individuals) attributable to different tastes/preferences and those attributable to different commodities.

Thirdly and finally, there is reasonable compelling evidence that suggests the possibility of resolving most, if not all, of the above-mentioned issues (as they relate to a large class, but not all, of environmental commodities by thoughtful design of the CVM -- considerable hueristic inquiry remains, of course, for identifying and verifying 'appropriate' designs which mitigate or eliminate above-described sources for bias. There remains an issue the substance of which is not related to questions of design, however, viz., the large body of evidence that supports the proposition that choices involving actual payments are substantively and significantly different from choices involving hypothetical payments. Given the relevance of the results from our review of advances made in Experimental Economics (Chapter IV) for an assessment of the implications of this issue, we defer further discussions to Chapter VI wherein results from all chapters are integrated to the end of offering tentative conclusions as to the state of the arts of the CVM.

ENDNOTES

Chapter V

1) Bishop and Heberlein, 1979, p. 327.

2) Bohm, 1972, p. 125. Interestingly, when individuals asked hypothetical questions and were then asked for actual payment, only 18 out of 54 changed their responses, an outcome interpreted by Bohm as reflecting people's reluctance to "... imply a confession that they had lied in the first round," p. 126.

3) As examples, T. Feather, 1959 and P. Slovic and S.C. Lichtenstein, 1968.

4) See, e.g., G.J. Stigler and G.S. Becker, 1977.

5) As examples, see T.C. Schelling, 1978; and J.W. Pratt, D. Wise and R. Zeckhauser, 1979.

6) Such an approach is seen in expressed efforts"... to determine if people will actually pay (as measured by a HPM measure) what they will pay (a hypothetical payment measured by the CVM)", in Schulze et al., 1981, p. 167.

7) See Burness et al., 1983, pp. 680-682 and Schulze et al., July, 1983, pp. 148-150.

8) Randall et al., 1983, p. 639.

9) Id, p. 646.

10) This is an argument made in Schulze et al., July, 1983, Chapter 1; and Burness et al., 1983.

11) See Schulze et al., July, 1983, section 1.F and Desvousges et al., 1983, Chapter 8.

12) See also Kunreuther with Ralph Ginsberg and Louis Miller, 1978. As another example of related results, see L. Robertson, 1974.

13) As an example, see D. Kahneman and A. Tversky, 1972; A. Tversky and and D. Kahneman, 1973; S. Lichtenstein, B. Fischhoff et al., 1978; and B. Fischhoff, 1975.

14) Kahneman and Tversky, 1979, Tversky and Kahneman, 1981, and P. Schoemaker, 1980.

15) See section 1.C in Schulze et al., 1983, for a discussion of experimental results suggestive of the mental account notion.

16) See also ad passim in Schulze et al., July, 1983a, p. 170; see also an earlier draft dated April, 1981)

17) See, for example, T.D. Crocker and R.G. Cummings, 1984. There is yet another functional relationship of potential importance, viz., "... the physical production and transformation linkages between public policies and (environmental/recreational) values", S.S. Batie and L. Shabman, 1979.

18) R.G. Cummings, H.S. Burness and R.D. Norton, 1981.

19) See, e.g., S. Gerking and W.D. Schulze, 1981.

20) See Cummings et al., 1983, and Schulze et al., July, 1983.

21) Particularly see Desvousges et al., 1983, and Schulze et al., July, 1983, (the Grand Canyon experiment).

22) As examples, see N. Georgeseu-Roegen 1958. N.E. Devletoglou, Feb., 1971 and R.D Luce, 1956.

23) We find a second CVM study involving unknown C(a) and reliance on C(p) in Burness et al., 1983, (also reported in Schulze et al., July, 1983). Unfortunately, the authors of this study did not examine the implications of varying C(p)'s on derived WTP.

24) As examples, see P.J.H. Schoemaker, and H.C. Kunreuther, 1973, pp. 603-18; J.C. Hershey and P.J.H. Schoemaker, 1980; R.S. Gregory, 1982; and R. Thaler, 1980.

25) See previously cited works by Schulze et al., July, 1883, and Burness et al., 1983. See also I. Majid, J.A. Sinden and A. Randall, 1983.

26) The context for the citation given above is "... variations in the materials describing the contingent market . .."; Ibid.

27) See J.D. Hey, 1983; W. Edwards, 1954; and, more generally, G. Stigler, 1950.

28) "The (only) way that uncertainty enters into the choice problem is when the choice must be made before it is known which ... (post-choice state of the world) . . . will prevail." Hey, 1983, p. 131.

29) An interpretation admittedly imputed to Randall's statement by the authors in their best efforts to understand the point argued in the statement.

30) Schoemaker (1982, pp. 545-547); see also K.E. Boulding, 1975, p. 84.

31) We note, however, the potential relevance of section V.B's discussion of hypothetical v. actual payment for the framing of WTP questions.

32) See also the use of 'budget constraint' arguments in assessing the time-preference research hypothesis (3') in section V.C.

33) Two of the three studies reviewed in Schulze et al., 1981, and studies by J.T. Daubert and R.A. Young, 1982, and D.A. Greenly, R.G. Walsh and R.A. Young, 1981.

34) In example, Schuman and Johnson, 1976; and I. Ajzen and M. Fishbein, 1977.

35) Grether and Plott, 1979, this consistency is noted by Randall et al., 1983. See also Pommerehne, Schneider and Zweifel, 1982; and Reilly, 1982.

36) For example, Burness et al., 1983, offer an EPA regulation on hazardous waste disposal as a commodity when, it would seem, individuals are valuing their perceptions of changes in risk.

VI. COMPARISON STUDIES: WHAT IS ACCURACY?

A. INTRODUCTION.

Thus far, we have examined results from studies involving experiments with the CVM, as well as from the psychology literature and studies from experimental economics, to the end of inquiring as to the extent to which potential sources for biases identified in Chapter II have been addressed in works accomplished to date. At this point, the litany of potential sources of bias in CVM measures, along with pro-con arguments relevant for each source presented above, may seem overwhelming; after reading these chapters, the reader may consider the case made for the psychologists' concern with problems associated with "limited capacity for information processing." In any case, one sees in these discussions the fundamental issue which must be faced if we are to meet the challenge of an objective assessment of the CVM; this issue is described by the question: against what criteria is the accuracy of the CVM to be evaluated? It would be inaccurate to say that scholars working with the CVM have ignored the issue of assessment criteria; it would be accurate to describe a large part of the efforts to address the issue as imprecise and intuitive. In looking to the CVM literature, the bulk of empirical evidence offered in these regards is seemingly limited to observations concerning the substance of CVM measures of the sort: 'this' evidence suggests that it's good, 'this' evidence suggests that it's bad. The inability to weight evidence had invited recourse to 'counting' types of assessments as a means for establishing accuracy in CVM measures. As examples in this regard, "(CVM studies) have generated a 'solid core' of value information which performs well ..." (Randall et al., 1983, p. 640); "More verification of (CVM) ... results through repeated application and comparison with actual behavior ... is necessary" (Rowe and Chestnut, 1983, p. 409); "There is no objective, a priori manner by which the accuracy of survey measures can be proven (or ... disproven ...); if successful, however, repeated experiments ... (may redefine) ... economists' reservations ... (about the CVM)." (Cummings et al., 1983, p. 12)

In considering the question as to appropriate criteria against which to assess the accuracy of measures derived by the CVM, two issues are of primary importance. First, it is useful to recall the rationale for our interest in the method. As discussed in detail in Chapter II, benefit-cost analysis is used, however imperfectly (sections II.B and II.C), in assessing optimal levels for a public investment. At a conceptual level, applications of benefit-cost analysis may be viewed as efforts to deduce market outcomes (vis-a-vis the level of public investment) that would obtain if such investments were made under market conditions. Given benefits (prices) and costs determined by market institutions, public investments would be provided at levels at which marginal benefits equal marginal costs.

Of course, for most pure public goods -- particularly environmental goods -- market institutions do not exist. The CVM is then used as a substitute for the 'missing' market; it is used to simulate the market in the sense of eliciting revelations of preferences (a willingness to pay) analogous to those which would have resulted under market conditions. Like the market institution, the CVM must then be viewed as an 'institution'. Thus, the general criterion against which to assess the CVM becomes clear: the extent to which the CVM institution, and preference revelations drawn therein, is comparable with the market institution and preference

revelations drawn therein.

The second issue of primary importance for our discussions concerns the notion of "accuracy" per se; i.e., what is (what do we mean by) "accuracy"? Notwithstanding the many potential sources of bias in CVM measures identified and discussed in earlier chapters, we must ultimately address the question: how accurate are values obtained from CVM studies? Are these values as accurate as values obtained from other traditional approaches such as the travel cost method (TCM) or the hedonic price method (HPM)? Obviously, if both the CVM and, for example, the HPM give the same value for the same commodity under the same circumstances and if this can be shown to be true when repeated for many environmental commodities, and, if the HPM is viewed as providing accurate measures of value, then this may provide strong evidence vis-a-vis the accuracy of CVM measures. Unfortunately, as we argue below, all of the comparison studies undertaken to date have failed to carefully assess the accuracy either of the CVM used *or* the accuracy of the HPM (or TCM) used for comparison. This lack of uniform approach for evaluating accuracy across the many individual comparison studies has led to confusion and inconsistency in interpreting the available evidence.

In efforts to address these issues, our discussions proceed as follows. In sections B and C we review results from the various studies which compare values derived from the CVM with values derived from alternative methods -- primarily the TCM and the HPM. In reviewing these studies, the implications of any study's results vis-a-vis the accuracy issue is considered within the limited context of statistical comparisons or, more often, less formal comparisons offered by the study's authors. In Section D we consider results from comparison studies within a broader context for "accuracy"; as a part of these latter discussions, we consider alternative, related, scientific definitions for the accuracy of measured values. In section E we examine the implications of scientific notions of accuracy, as they are used in weighting the results from comparative studies, for means by which the CVM might be assessed in state-of-the-arts terms. Concluding remarks are offered in section F.

B. VALUE COMPARISONS: THE CVM AND THE TCM.

Five major studies have been completed wherein primary attention is given to the comparison of non-market values for environmental commodities derived via the CVM with those derived from the travel cost method (TCM). These are the studies reported by Knetsch and Davis; Bishop and Heberlein; Desvousges, Smith, and McGivney; Thayer; Seller, Stoll, and Chavas; and Fisher.

1. Knetsch and Davis. The earliest study comparing value estimates obtained from the CVM with estimates derived from other procedures is reported by Knetsch and Davis (1966). The authors compared three methods of measuring the benefits of recreation in the woods of northern Maine. Using data obtained from an earlier survey by Davis (1963) they compare willingness-to-pay estimates resulting from an application of the CVM to values related to individuals' 'willingness to drive' and to values derived from the TCM.

CVM interviews were conducted in the Pittson Farm area (in northwestern Maine) of 185 hunters, fishers and campers using the area. The respondents were asked if their decision to use the site would change if the cost of doing so increased. Costs were then systematically increased until the respondent switched from 'inclusion in' to 'exclusion from' the activity. For respondents who thought the original amount excessive, costs were decreased until they switched from 'exclusion from' to 'inclusion in' the recreation activity. The final amount was used as their maximum willingness-to-pay to participate in recreation activities at the Pittson Farm area. The mean willingness-to-pay was $1.71 per household per day; obtained values ranged from zero to $16.66.

A measure of willingness-to-pay was then derived by a multiple regression analysis of data derived via the CVM which demonstrated that nearly sixty percent of the variance in bid values could be explained by differences in household incomes, degree of familiarity with the site (Note: perceptions of the 'commodity'?) and the average length of each visit. By administering a questionnaire to users stopped at a traffic checking station, estimates of income, length of stay and degree of site familiarity for the user population were obtained. With these two pieces of information, a demand schedule and total recreation benefits were estimated. The demand schedule was derived from ordering the user population by calculated willingness-to-pay, and the benefits were computed from the area under the demand schedule from the highest price to the price considered. Their estimatae of maximum benefits (when 'price' is zero) to the 10,333 household days of recreation translates to a WTP of $1.71 per household per day.

Knetsch and Davis then develop two additional estimates of willingness-to-pay. The first estimate is based on 'willingness to drive' (WTD), a method earlier proposed by Ullman and Volk (1961). Individuals, the same individuals interviewed for the CVM, were asked how much further (in miles, beyond the Pittson area) the individual would drive to avail himself/herself of recreation facilities like those in the Pittson area if they were no longer to have access to the Pittson area. The authors assert that "... willingness-to-pay was found to increase about five cents per mile as a function of willingness-to-drive additional miles" (Knetsch and Davis, 1966, p. 137). A development of this finding is not given in the paper. Using

this 5 cents/mile, WTD data are used to estimate benefits attributable to the Pittson recreation area; estimated maximum benefits, the area under the derived demand curve at a zero 'price', were $64,000, which compares with $72,000 derived via the CVM.

The second alternative (to the CVM) value derived by Knetsch and Davis was estimated with the TCM. Visitation rates of visitors from groups Of counties were plotted against travel distance. The resulting 'visitor days as a function of distance travelled' relationship was then converted into a 'visitor days as a function of costs' via costing distance at 5 cents per mile for one-way distance; travel costs for 1,327 respondents (out Of a total population of 6,678) for whom Pittson was not the primary destination of their trip were arbitrarily weighted at .5. These TCM procedures yielded an estimate of maximum benefits, as defined above, in the amount of $70,000.

Knetsch and Davis acknowledge the crudeness of approximations derived in their WTD and TCM estimates, a topic which we will not consider here (Mendelsohn and Brown, 1983); of interest here are Knetsch and Davis's value comparisons. Knetsch and Davis do not subject their CVM, WTD, and TCM benefit estimates to statistical analysis in comparing them. Rather, their discussions in these regards focus simply on the demonstrated 'closeness' of their results: i.e., upon casual inspection, $72,000 (benefits based on the CVM), $64,000 (benefits based on the WTD method) and $70,000 (benefits based on the TCM) are 'close'. Given the sharp divergence and disparities in assumptions underlying the three measures, the 12% maximum difference between the measures is indeed remarkable. Little basis exists, however, for interpreting this 'closeness' beyond, perhaps, the authors' above-cited observation that such closeness may indicate some promise of the methods as a means for estimating benefits for recreation.

2. Bishop and Heberlein. The primary purpose of the paper by Bishop and Heberlein (1979) (hereafter, B-H) was to point out the biases that may result from the use of indirect and direct measures of values for non-market goods, specifically the TCM and CVM. After discussing several potential sources of bias with the techniques, they undertake an experiment designed to see how serious these biases actually might be.

B-H conducted three surveys of hunters who had received free early season goose hunting permits in 1978. Hunters were divided into three groups. The first sample of 237 received a cash offer in the mail for their permits. The checks ranged from $1 to $200, and the respondents were requested to return either the check or the permits. The second sample of 353 persons received a questionnaire by mail designed to elicit either their hypothetical willingness-to-sell their permit or their hypothetical willingness-to-pay for their permit. The third sample of 300 received a questionnaire designed to elicit factual information necessary to estimate a travel cost demand curve. The authors report a response rate of at least 80% for the three surveys, and report that the results of a comparison Of differences in socioeconomic and other characteristics found the three samples to be relatively homogeneous.

Results reported for the B-H study are given in Table 6.1. The actual cash offers resulted in a willingness-to-sell figure of $63 per permit. B-H note, however, that this figure may be conservative due to the $200 upper limit on offers; regression results indicated that 10% to 12% of these surveyed would have sold at a higher amount.

Table 6.1

Summary of Results[a]

Sample Group	Total Consumer Surplus	Surplus Per Permit
	(1978$)	
1. Actual Cash Offers	$ 880,000	$ 63
2. Hypothetical Offers:		
Willingness To Sell	1,411,000	101
Willingness To Pay	293,000	21
3. Travel Cost Estimates		
Model 1 (time value=0)	159,000	11
Model 2 (time value=1/4 median income rate)	387,000	28
Model 3 (time value=1/2 median income rate)	636,000	45

a. Source: Bishop and Heberlein (1979), p. 929.

The hypothetical willingness-to-sell figure was quite a bit larger: $101 per permit. Here too, the maximum offer of $200 created some difficulty. Regression results indicated that 35% of the hunters in this group would have (hypothetically) 'sold' if the offer were over $200. As a result, B-H assert that"... had the models been truncated at a higher figure the difference between willingness-to-sell measured using actual money and measured using hypothetical dollars would have been even more pronounced". (Bishop and Heberlein, 1979, p. 924) Their second comparison was between actual willingness-to-sell, hypothetical willingness-to-sell and hypothetical willingness-to-pay. Using the former as a measure of consumer surplus, (CS), they note, citing Willig (1976), that WTS > CS > WTP. However, B-H argue that "... for the range of values we are discussing here ($1 - $200) ... willingness-to-pay, and willingness-to-accept-compensation should be quite close together". (p. 929) This however, was not the result obtained by B-H. B-H report a WTP figure of $21 per permit, far below the $63 estimate of consumer surplus. Estimates of WTS and WTP, derived via the CVM, are then compared by B-H with three estimates of travel-costs, differing only in the valuation of time spent traveling. Following Cesario's (1976) suggestion that time be valued at between 1/4 and 1/2 the wage rate, B-H set up three different travel-cost models. The first does not include a value for time; the second model values time at 1/4 of median income and the third at 1/2 of median-income.

As Table 6.1 demonstrates, even when the time spent traveling is valued at 1/2 of median income, the travel cost estimate of $45 is substantively (29%) below the CVM estimate of $101; both TCM and CVM values differ substantively from the 'actual' cash offer ($63). Because of the divergence between the various measures tested, B-H assert that 'the results summarized here must be interpreted as supporting the hypothesis that the sources of bias listed above do have significant impacts on (CVM) and (TCM) values for recreation and other extra-market goods.' (p. 929)

As in the Knetsch and Davis study, B-H's comparisons of CVM measures with non-hypothetical (actual cash offer) measures and TCM measures is qualitative in nature; their a priori expectations for comparisons are that the measures "... should be quite close together" (p. 929) and data comparisons are analyzed in terms of percentage differences: "... the (TCM) estimate averages only $45.00, 29% below the (actual cash offer) benchmark figure of $63.00." (p. 929) We may then conclude little more than that, while Knetsch and Davis report CVM and TCM measures which are 'close', B-H report CVM and TCM measures which are not 'close'.

3. Desvousges, Smith and McGivney. The study by Desvousges, Smith and McGivney (1983) (hereafter, DSM) is of particular interest for our discussions of comparative values for several reasons. It is a recent study and the authors attempt to deal with many of the measurement/comparison problems encountered in earlier studies. Most importantly, the authors attempt to go beyond qualitative comparisons of CVM and TCM values in forming and testing hypotheses concerning the relationships between such values.

DSM make pairwise comparisons of the results from three different techniques for estimating benefits attributable to water quality improvements. The authors compare user values obtained from both the TCM and CVM, and option prices obtained from both the CVM and contingent ranking approaches. The commodities at issue in this study are water quality changes in the Monongahela River in Pennsylvania. Three different types of water quality changes were considered. The first was a decline in water quality resulting in a complete loss of recreational activity in the River. The second and third were increases in water quality from boatable to fishable and boatable to swimmable levels, respectively.

The authors surveyed 303 households in a five county region in Southwestern Pennsylvania, near the Monongahela River. Personal interviews were conducted from November through December 1981. As a part of the CVM, respondents had described to them the hypothetical market, the commodity to be valued and the payment vehicle (higher taxes and prices). Respondents were then asked their valuation of the commodity. A water quality ladder was used to help the respondent establish a linkage between an index of water quality and an associated recreation activity. The respondents were divided into four approximately equal sub-groupings. One group was given a payment card with values ranging from $0 to $775 in $25 increments, and were asked to pick any amount on the card, any amount in between the values listed, or any other amount. A second group was asked their valuation directly, without the use of a payment card or suggested starting point. The third and fourth groups were given a 'starting point', i.e., they were asked if they would be willing to pay $25 or $125, respectively. After their yes or no response, a bidding process was used until a maximum bid was obtained. Each group of respondents was asked their willingness-to-pay for three water quality changes: to avoid a decrease in water quality to the

point where the river could not be used; to raise the water quality level from boating to fishing quality; and to raise the level from boating to swimming quality. Those who gave a positive response to the boatable-fishable increment were asked their additional WTP to go from fishable to swimmable, Only those who gave a zero bid for the boatable-fishable increment were asked the boatable-swimmable question directly. For others, it was derived by adding boat-fish bids to fish-swim bids. After the final value for each of the changes was obtained, the respondents were asked how much of this value was attributable to their actual use of the River, a 'user value', and how much was attributable to their desire to maintain options for future uses, i.e., their 'option value'.

Finally, the survey respondents were asked to undertake a contingent ranking of options. They were shown four cards, on each of which was a water quality ladder with an annual payment amount of either $5, $50, $100, or $175 paired to no recreation, boatable, fishable, or swimmable recreation water quality levels, respectively. Respondents were asked to rank the combinations from most to least preferred. An ordered logit and an ordered normal procedure (see Rae, 1983) were used to estimate willingness-to-pay from the contingent ranking results.

DSM also used a generalized travel cost model to estimate recreation benefits. The model was developed from data drawn from 43 water-based recreation areas surveyed in the 1977 National Outdoor Recreation survey. The TCM data provided information on time spent at a given site, number of visits to the site, travel time to the site, and respondents' annual income. To measure travel cost, the distance to a given site was obtained from a Rand McNally Road Atlas. The marginal cost of driving to the site was assumed to be $0.08 per mile. Thus, travel costs were derived by multiplying the length of the trip (round trip miles) by mileage costs at $.08 per mile. Since hourly wages were not available in their data set, DSM used a semi-log hedonic wage model to estimate hourly wages for each individual in the sample. The mean estimated wage rate of $5.44 per hour was used as the opportunity cost of travel time, and onsite time. Of course, this method differs from the approach used by Bishop and Heberlein (1979) who, as noted above, valued travel time (only) from zero to 1/2 the wage rate.

The results of DSM's estimations of contingent valuation, contingent ranking and travel cost measures of water quality values are shown in Table 6.2 for each of the proposed water quality changes. Referring to Table 6.2, for increases in water quality from boatable to swimmable levels, the option prices obtained by the CVM range from about $25 to $60, depending on the valuation format used. Similarly, user values range from about $10.50 to $51.00 (users only, see footnote a). The Contingent Ranking Method (CRM) is used for estimating option prices only. Depending upon the statistical estimation technique used, the option price for the third category of water quality change was either $108 (ordered logit method) or $112 (ordered normal method). Similarly, the travel cost method yields but one value, the user value, which is about $15.00 for improvements from boatable to swimmable water quality.

Our interest is in DSM's analysis concerning value comparisons. In this regard, DSM compare the CVM with the TCM, and the CVM with the CRM. These comparisons involved two tests: a simple comparison of sample means, and a statistical comparison of individual values. In terms of CVM-TCM

TABLE 6.2 COMPARISON OF BENEFIT ESTIMATES FOR WATER QUALITY IMPROVEMENTS (1981 Dollars)

Methodology	ΔWQ = Loss of use			ΔWQ = Boatable to fishable			ΔWQ = Boatable to swimmable		
	Option price	User value[a]	Option value	Option price	User value[a]	Option value	Option price	User value[a]	Option value
I. Contingent valuation[b]									
Direct question	24.55	6.57 (19.71)	17.98	17.65	7.06 (21.18)	10.59	31.20	13.61 (31.18)	20.80
Payment card	51.00	6.20 (19.71)	44.82	29.26	9.72 (30.88)	19.54	42.87	15.92 (51.18)	26.76
Iterative bidding ($25)	28.97	2.16 (6.50)	26.81	15.95	1.38 (4.21)	14.57	25.09	3.12 (10.53)	21.64
Iterative bidding ($125)	57.40	12.08 (36.25)	45.31	36.88	6.77 (20.31)	30.10	60.20	13.43 (48.75)	43.96
II. Contingent ranking[c]									
Ordered logit	-	-	-	60.03	-	-	108.06	-	-
Ordered normal	-	-	-	62.12	-	-	111.81	-	-
III. Generalized travel cost[d]	-	82.65	-	-	7.01	-	-	14.71	-

[a]The numbers in parentheses below the estimated user values report average values for users only. Since nonusers have a zero user value, the combined mean understates user values.

[b]These estimates are for the combined sample including users and nonusers. It excludes protest bids and outliers detected using the Belsley, Kuh, and Welsch regression diagnostics.

[c]These estimates are for the sample of respondents with usable ranks and reported family income. Estimates evaluated at the intermediate payment level.

[d]These estimates are for survey respondents using Monongahela sites and have been converted to 1981 dollars using the consumer price index.

Source: Desvousges, Smith and McGivney, 1983, pp. 8-13.

comparisons, the first test, a simple (i.e., non-statistical) comparison of means tested the hypothesis that the CV bid would be less than the TC measure for water quality improvements, with the difference being slight, about 5%. Thus, they test H_o: CV = .95TC. For water quality improvements, CV is greater than TC, except for the $25 format, where CV < .95TC. (See Table 6.2). In the case of a loss in water quality, CV is less than TC, as expected, but much less than .95TC; the TC estimate is more than two times larger than the CV measure. The authors argue that this large disparity was likely the result of failure to consider the effect of substitute sites as an argument in the demand function for a particular site, overestimation of the TC measure of ordinary consumer surplus for loss in water quality. In spite of this, the authors express some surprise at the difference in magnitudes and directions of differences between TC and CV estimates.

But these were not statistical tests. Furthermore, the relevant comparison, they argue, is against individual benefit measures. To make these comparisons, they regress the CV measure of user value on the TC measure, using dummy variables for three of the bid elicitation methods. In this respect, they test three hypotheses. If, as theory predicts, the CV measure is only slighty smaller than the TC estimate, then the intercept of the OLS equation should not be different from zero. Equally important, if the two methods result in comparable values, then the coefficient on the TC measure should not be different from unity. If the valuation method used in the CV survey has no influence on the resulting bid, then the coefficients on these variables should not be different from zero.

The results of these tests are shown in Table 6.3. As in their 'simple' tests, the relationship between CVM and TCM values differs in the quality-loss case from that in the quality-improvement cases. In the case of a loss in water quality, their test results seem somewhat ambiguous. The test fails to reject the hypothesis of zero intercept, suggesting that the CV and TC measures are similar. But the test for unitary slope (see footnote b in Table 6.3) rejects the hypothesis, suggesting that, given the magnitude of the coefficient on TC, CV measures are much less than TC measures of user values. The reason for the disparity, they argue, seems to lie in the overstated TC estimates (mentioned above). "Based on the association between estimates across individuals, there is support for the conclusion that the travel cost model overstates the benefits associated with avoiding the loss of the area." (Desvousges, Smith, and McGivney, 1983, p. 8-17) Thus the statistical test results seem to support the conclusion of the 'simple' test.

In both cases involving water quality improvements their test results are clearer. Both the null hypothesis of zero intercept and unitary slope (see footnote b, in Table 6.3), are rejected at the 10% level. Since both tests agree, the results strongly indicate no association between the TC and CV estimates. The authors, however, caution against so strong an interpretation of the results, because "the generalized TC model does not permit the effect of the intercept to be distinguished from at least one of the questioning formats. In the models reported in Table 6.3, the intercept reflects the effects of the iterative bidding format with a $125 starting point." (p. 8-17) They also note that "... there is some (ambiguous) evidence to support the conclusion that contingent valuation method may overstate willingness-to-pay for water quality improvements". (p. 8-17) DSM's conclusions do not effectively speak to the ambiguities that arise from the

TABLE 6.3 A COMPARISON OF CONTINGENT VALUATION AND GENERALIZED TRAVEL COST BENEFIT ESTIMATES[a]

	ΔWQ = Loss of area		ΔWQ = Boatable to fishable		ΔWQ = Boatable to swimmable	
	Model	Test[b]	Model	Test[b]	Model	Test[b]
Independent variable						
Intercept	21.862 (1.371)	-	33.985 (1.900)	-	59.574 (2.017)	-
Travel cost benefit estimate	.328 (1.169)	-4.357	-3.670 (-1.204)	-1.712	-2.713 (-1.141)	-1.793
Qualitative variables						
Payment care	-32.640 (-2.551)	-	51.757 (2.639)	-	77.010 (2.359)	-
Direct question	-14.602 (-1.270)	-	12.957 (0.748)	-	21.001 (0.729)	-
Iterative bid ($25)	-31.817 (-2.549)	-	-11.244 (-0.595)	-	-21.819 (-0.693)	-
R^2	.099		.120		.107	
n	93		93		93	
F	2.42 (0.05)[c]		3.00 (0.02)[c]		2.62 (0.04)[c]	

[a]The numbers in parentheses below the estimated coefficients are t-ratios for the null hypothesis of no association.

[b]This column reports the r-ratio for the hypothesis that the coefficient for the travel cost variable was 1.55. The travel cost model measures consumer surplus in 1977 dollars. The contingent valuation experiments were conducted in 1981. Using the consumer price index to adjust the travel cost benefit estimates to 1981 dollars would require multiplying each estimate by 1.55. Since the estimated regression coefficients (and standard errors) will correspondingly adjust to reflect this scale change, a test of the null hypothesis that the coefficient of travel cost was equal to unity is equivalent to a test that is equal to 1.55 when the travel cost benefit estimates are measured in 1977 dollars and user values estimates (the dependent variable) are in 1981 dollars.

[c]This number in parentheses below the reported F-statistic is the level of significance for rejection of the null hypothesis of no association between the dependent and independent variables.

Source: Desvousges, Smith and McGivney, 1983, pp. 8-16.

stark differences in CVM-TCM relationships seen in the quality-loss and quality-improvement contexts. These differences invite speculation as to the relevance of 'threshhold' effects (Crocker, 1984) for their analysis, particularly in light of the positive relationship between CVM and TCM measures in the quality-loss case and negative relationships between the two measures indicated in the quality-improvements cases.

In terms of comparing CVM measures with those derived via the Contingent Ranking (CR) method, both methods undertake to measure compensating surplus, thus the null hypothesis tested is that CV = CR. As Table 6.2 reveals, however, the CR approach results in values that seem consistently higher than CV values for water quality improvements. To test the statistical significance of these differences, DSM regress the CV measure of option price on the CR measure, again using dummy variables for three of the bid elicitation modes, for improvements in water quality -- CR measures were not obtained for the water quality-loss case. Since the CR value depends upon the payment level suggested by the cards presented to the respondent, regressions were run for each of three different payment levels; $50, $100, and $175. The results are shown in Table 6.4. As noted above, two econometric estimating techniques were used, ordered logit and ordered normal. The three statistical hypotheses for these regressions are the same as noted above. In this case, however, neither the hypothesis of zero intercept nor Of unitary slope (Test Column) can be rejected at the 90% level. This results in the failure to reject the hypothesis that CV = CR: thus, the contingent valuation and ranking techniques move in the same direction across individuals, with the CR estimates not significantly different from the CV estimates. The authors warn ,however, that despite the fact that both methods attempt to measure option price, since the same survey asked for CV and CR estimates, the strong relationship between them may simply reflect the respondent's efforts to appear consistent.

In summary, DSM's value comparisons between the CVM and TCM and between the CVM and CRM yield interesting, but somewhat ambiguous results. The authors find CV measures to overstate WTP for improvements in water quality as compared to values measured by the TCM. Curiously, however, they argue that these differences "... are not substantial and fall within the range of variation of the contingent valuation estimates across the question formats." (p. 8-21) In spite of the ambiguity of the test results, the authors argue that, for losses in water quality, the CV measure is found to be roughly consonant with the TC measure. The authors do find unambiguous close agreement between the CV and CR measures of WTP.

4. Seller, Stoll and Chavas. One of the more recent study comparing travel cost and contingent survey methods is by Seller, Stoll and Chavas (1984) (hereafter, SSC). The authors compare a regional TCM with two forms of the CVM: an open-ended questionnaire format (similar to DSM's direct question approach) and a close-ended format (multiple starting points). Since the authors assert that the reference level of utility is nonparticipation in the activity, an equivalent measure of willingness-to-pay is derived.

The interviews were conducted with past and present users of one of four lakes in Eastern Texas: Lakes Conroe, Livingston, Somerville, and Houston. The authors used a mail questionnaire to gather the travel cost and contingent valuation data. The questionnaires were mailed to 2000 registered boat owners in the 23 county area surrounding the four lakes, identified as the major origin of most users.

TABLE 6.4 A COMPARISON OF CONTINGENT VALUATION AND CONTINGENT RANKING BENEFIT ESTIMATES

	ΔWQ = Boatable to fishable						ΔWQ = Boatable to swimmable					
	Payment = $50		Payment = $100		Payment = $175		Payment = $50		Payment = $100		Payment = $175	
Independent variable	Model	Test	Model	Test	Model	Test	Model	Test	Model	Test	Model	Test
ORDERED LOGIT												
Intercept	-20.141 (-1.095)	-	-23.647 (-1.223)	-	-23.927 (-1.227)	-	-25.661 (-0.795)	-	-30.734 (-0.905)	-	-31.032 (-0.906)	-
Δ Payment	1.209 (4.279)	0.741	1.315 (4.237)	1.016	1.330 (4.214)	1.048	1.081 (3.925)	0.293	1.170 (3.867)	0.561	1.183 (3.841)	0.594
Qualitative variables												
Payment card	-22.486 (-2.424)	-	-22.070 (-2.380)	-	-21.960 (-2.367)	-	-46.842 (-2.877)	-	-46.145 (-2.834)	-	-45.961 (-2.822)	-
Direct question	-35.267 (-3.751)	-	-34.595 (-3.683)	-	-34.425 (-3.665)	-	-55.327 (-3.353)	-	-54.215 (-3.288)	-	-53.935 (-3.270)	-
Iterative bidding ($25)	-38.045 (-4.067)	-	-37.562 (-4.015)	-	-37.446 (-4.001)	-	-68.611 (-4.178)	-	-67.817 (-4.128)	-	-67.626 (-4.115)	-
R^2	.165		.164		.163		.153		.151		.150	
n	184		184		184		184		184		184	
F	8.87 (0.0001)		8.77 (0.0001)		8.72 (0.0001)		8.06 (0.0001)		7.94 (0.0001)		7.88 (0.0001)	
ORDERED NORMAL												
Intercept	-13.467 (-0.839)	-	-15.565 (-0.940)	-	-15.832 (-0.951)	-	-15.153 (-0.537)	-	-18.212 (-0.626)	-	-18.559 (-0.634)	-
Δ Payment	1.073 (4.554)	0.309	1.140 (4.528)	0.554	1.151 (4.516)	0.592	.962 (4.182)	-0.165	1.018 (4.146)	0.073	1.028 (4.131)	0.113
Qualitative variables												
Payment card	-22.642 (-2.457)	-	-22.357 (-2.426)	-	-22.286 (-2.418)	-	-47.108 (-2.910)	-	-46.630 (-2.880)	-	-46.510 (-2.872)	-
Direct question	-34.934 (-3.745)	-	-34.458 (-3.696)	-	-34.344 (-3.683)	-	-54.808 (-3.345)	-	-54.020 (-3.298)	-	-53.832 (-3.286)	-
Iterative bidding ($25)	-37.541 (-4.014)	-	-37.196 (-4.004)	-	-37.116 (-3.994)	-	-67.808 (-4.156)	-	-67.242 (-4.120)	-	-67.112 (-4.111)	-
R^2	.176		.175		.174		.162		.160		.160	
n	184[b]		184[b]		184[b]		184[b]		184[b]		184[b]	
F	9.53 (0.0001)[b]		9.47 (0.0001)		9.43 (0.0001)		8.63 (0.0001)		8.54 (0.0001)		8.51 (0.0001)	

[a]These estimates are for the combined sample including users and nonusers. It excludes protest bids and outliers detected using the Kuh-Welsch regression diagnostics.

[b]These estimates are for the sample of respondents with usable ranks and reported family income.

Source: Desvousges, Smith and McGivney, pp. 8-19.

The TCM involved estimating a system of demand equations,

$$V_{ij} = \alpha_j + \sum_{k=1}^{4} \beta_{jk} C_{ik} + \delta_j Y_i + \gamma_j Z_i + \epsilon_{ij} \qquad (1)$$

where

V_{ij} = the number of visits to the jth site (j = 1...4) by the ith household,

C_{ik} = costs incurred by household i while at and traveling to site k (k = 1...4).

Y_i = income of household i

Z_i = demographic variables,

$\alpha_j, \beta_j, \delta_j$ = parameters to be estimated, and

ϵ_{ij} = error term.

Costs were measured as gasoline expenses only, with the value of travel time set at zero, using the equation

$$C_{ik} = (2d_{ik}/mpg_i \times 1.10) + E_{ik} + (gas_{ik} \times 1.10) + fees_{ik},$$

where

d_{ik} = one-way distance for household i traveling to site k,

mpg_i = average miles per gallon on household i's vehicle,

1.10 = average cost of gasoline (1980 dollars per gallon),

E_{ik} = other variable costs incurred by household i traveling to site k,

gas_{ik} = number of gallons of gasoline used by the pleasure boat,

user and/or entrance fees.

Specifying a priori a linear system of equations, benefits from each site were measured using the TCM as

$$M = \int_{C_j}^{C_j} V(C_j...)dC_j, \qquad (2)$$

where M = Marshallian consumer surplus

dC_j = change in travel costs, with C_j the vertical intercept on V.

Of the 2000 questionnaires mailed out, 731 were used to gather travel cost data. The four demand curves generated from the data using equation 1, holding Y and Z constant, are shown in Table 6.5. The authors do not report standard errors or t-statistics associated with the coefficients. The average (Marshallian) consumer surplus associated with each site was calculated as the area under V above the current expenditure level at the mean number of visits for each lake. The results are shown in column 3 of Table 6.5. As is apparent by the results, willingness-to-pay for recreation at the Lake Livingston greatly exceeds that for the other three areas combined.

Table 6.5

Results of the TCM

Area	Demand Equation[a]	Average Consumer Surplus
Lake Conroe	$V_1 = 14.46 - 0.23C_1$	$32.06
Lake Livingston	$V_2 = 10.04 - 0.12C_2$	$102.09
Lake Somerville	$V_3 = 8.63 - 0.13C_3$	$24.42
Lake Houston	$V_4 = 3.28 - 0.04C_4$	$13.07

a/ V_j = number of visits at site j, (j = 1...4) and C_j = cost of visiting site j.

The CVM used two different bid elicitation approaches. One was an "open-ended" approach wherein the respondent specifies the initial value of the bid, a direct question approach similar to that used by Desvousges, Smith, and McGivney. The other was a "close-ended" approach wherein the respondent is given an "estimate" of the cost and asked to respond "yes" or "no" to the willingness to pay question.

Respondents to both forms of the survey were asked their willingness-to-pay an annual fee for a boat ramp permit. Two questions were asked in the open-ended format:

(1) How high could costs go to keep you using this site just as often?; and

(2) How high could costs go if you were restricted to using this site half as often.

Answers to these questions were used as two points on a Bradford-type bid curve for each individual. The bid curve is specified as

$$WTP = F(Q,Y)$$

where

WTP = the Hicksian equivalent measure of willingness to pay,

Q = the number of visits to the site (annually), and

Y = the respondents' income.

Of the 2000 questionnaires sent out, 275 using the open-ended format were used. The bid curve was estimated from this data using three different functional forms: linear, linear with a squared term in Q, and double logarithmic. The authors differentiated the log form of the bid curve to find the inverse Hicksian demand curve. Since the reference level of utility is nonparticipation in the recreation activity, the area under the Hicksian demand curve at the mean number of visits is the equivalent measure of consumer surplus. The demand curves and surplus measures are shown in Table 6.6.

Table 6.6

Results of the Open-Ended CVM

Area[a]	Demand Equations	Surplus Gross	Surplus Net[b]
Lake Conroe	$dWTP/dV = 1.79V^{-.75}$	$9.06	-$8.65
Lake Livingston	$dWTP/dV = 1.52V^{-.80}$	$8.87	$1.09
Lake Houston	$dWTP/dV = 1.22V^{-.70}$	$3.81	-$2.28

a The results of the demand relationship for Lake Somerville were considered by the authors to be unreliable because the demand curve was not downward sloping and lay in the fourth quadrant. Hence no results for Somerville were reported.

b Net surplus values were obtained by subtracting average launch fee expenditures from gross surplus.

Reflecting on the negative values for the surplus measures at Lakes Conroe and Houston, the authors conclude:

> "The negative values ... seem to indicate that people reported they were willing to pay less for an annual ramp permit than they already paid in total launch fees over the year on a per visit basis." (p. 22)

They argue that the negative and low results indicate that the open-ended questionnaire technique may be unreliable.

For the close-ended format, respondents were asked to respond "yes" or "no" to the following question:

> "If the annual boat ramp permit cost $X in 1980, would you have purchased the permit so that you could have continued to use the lake throughout the year?" (p. 15)

Ten values for $X were used, ranging from $5 to $300. The authors use a binary response model (because the answers are binary -- yes or no) to analyze the results. Assuming a logistical cummulative distribution function, a logit procedure (using maximum likelihood estimation) was used to estimate the probability that the respondent will answer "no" to a given value of X.

Varying the number of annual visits from 1 to 30, a Bradford-type curve was derived for each of the lakes. Of the surveys mailed out using the close-ended format, 211 were used. Differentiating the bid curves-produced

a Hicksian demand curve for each lake. Finally, the area under each demand curve at the mean number of visits to each lake is the gross measure of willingness-to-pay. The results are shown in Table 6.7.

Table 6.7

Results of the Close-Ended CVM

Area[a]	Gross Surplus	Net Surplus[b]
Lake Conroe	$53.94	$39.38
Lake Livingston	$42.40	$35.21
Lake Houston	$36.34	$31.81

a Again, the results from Lake Somerville fail to produce negatively sloped demand curves, hence were considered unreliable.

b Net surplus values were obtained by subtracting average launch fees from gross surplus.

The authors compare the results of the TCM with both CVM formats, with two caveats in mind. First, the TCM produces a Marshallian measure of consumer surplus, while the CVM produces a Hicksian measure of equivalent variation. However, since the authors report a small income effect they note that the difference should be small. Second, they note that the TCM produces results for boating only. Thus, they assert that this may cause a small divergence in the two measures. The hypothesis tested in the comparison is that the CVM value will exceed the TCM value: CVM > TCM (Although they state the difference to be small, the authors do not specify how small, only "comparable").

Confidence intervals are established at the 95% level to test for similarity in the bids. The results of the tests are reported in Table 6.8. For the open-ended questions, the null hypothesis of "comparable" means was rejected at each of the sites. As is clear in Table 6.8, the open-ended questions consistently produce smaller (in some cases negative) estimates of average consumer surplus. For the close-ended questions the null hypothesis is not rejected, the mean bids derived from the PCM and CVM are statistically equal.

In summary, one comment is in order. SSC attempt to determine the accuracy of the reported bids by relying on respondents' assessment of the accuracy of their stated bid. Survey participants were asked if they felt their stated willingness-to-pay to be "quite accurate", "accurate in a ball park kind of way", or "there is no way I could come up with accurate answers". They report that the majority (63.4%) of the respondents to the close-ended questions felt their bids were "quite accurate", while the (41%) of the respondents to the open-ended questionnaire felt they could only give "ball park" accurate responses. In addition, they report that the portion of "inaccurate" responses was higher for the open-ended format (24.8%) than for the close-ended format (9.2%). However, it seems fair to say that one can not, in fact, conclude that the close-ended question format produces results which are more reliable than alternative formats. In addition, a

TABLE 6.8 LOGIT ANALYSIS OF THE CLOSE-ENDED FORM OF THE CONTINGENT VALUATION METHOD

Lake	Estimated Coefficients (t Statistics) Intercept (lnα)	Suggested Price	Number of Visits	ρ^2 [a]	N	Percent of Correct Forecasts
Conroe	-6.13*** (-2.88)	1.79*** (3.53)	-.16 (-.47)	.39	70	87
Livingston	-3.06** (-1.86)	1.37*** (2.92)	-.67** (-1.75)	.39	74	85
Somerville	-4.78*** (-2.48)	1.26*** (2.92)	.88 (1.54)	.40	47	87
Houston	-2.32 (-.88)	.99 (1.75)	-.47 (-.84)	.31	15	80

[a] ρ^2 = goodness-of-fit (analogous to R^2)

*** Significant at the .01 level of confidence

** Significant at the .05 level of confidence

one-shot response to a single yes-no question gives much less information than someone's open-ended direct response; e.g., even if a response of $75 is fairly inaccurate, it probably tells us more than if the respondent said "yes" to the question "would you be willing to pay $10?" Slovic, et al. (1980) as well as Kahneman and Tversky (1974) report that individuals are consistently observed to overstate the degree to which their responses to questions involving some uncertainty are accurate. (See Chapter V for a discussion of this issue.)

5. Thayer. Thayer's (1981) study involves the comparison of values derived via the CVM with values derived from a variant of the TCM, as the TCM is generally structured. Thayer compares CV values with values derived from a 'site substitution' method (SSM) which, as will be shown, is reminiscent of Knetsch & Davis' 'willingness to drive' method.

Thayer's concern is in comparing CVM values with values from the SSM as well as in testing methods for dealing with starting point, hypothetical and information biases -- biases which are discussed above in Chapter III. Thayer conducted a survey in the Jemez Mountains of northern New Mexico. Recreators in the area were asked their willingness to pay an entrance fee to prevent the development of a geothermal power plant in the Jemez Mountains. They were also queried as to contingent site substitution plans should the plant ultimately be constructed.

Respondents were shown photographs of geothermal developments in other wilderness sites, and a map of the area where the Jemez plant would be built. In addition, the increased noise level and odors associated with geothermal power plants were described in detail. A bidding procedure was then initiated, following closely the methods used in Randall, et al. (1974).

Thayer attempted to control for starting point bias by separating the respondents into two groups. For the first group, bids began at $1 and were increased in whole dollar increments until the respondent would pay no more, whereupon the amount was decreased in quarter dollar decrements until a 'no more' response was given. For the second group, the bidding process was reversed, bids began at $10, were decreased in dollar amounts, then increased in quarter amounts. A comparison of the mean bids from the first group with the second group showed the bids to be not significantly different at the 10% level.

The final test was for hypothetical bias. It was in this regard that Thayer compared results from the CVM with those from the SSM. His hypothesis was that cost of traveling to a substitute recreational area represented a minimum loss in consumer's welfare from development in the Jemez. Thus, site substitution costs should represent at least the minimum they would be willing to pay to prevent development of the geothermal power plant. If the site substitution measures are similar to derived CV values, he argues, then CV values are not influenced by the hypothetical nature of CVM.

Due to data limitations, Thayer was unable to perform a comparison-of-means test. Thus, as in most earlier studies, his value comparisons are qualitative in nature. Thayer observes that the range of values for additional SS travel costs -- from $1.85 to $2.59 -- brackets the mean willingness-to-pay estimate from the CVM of $2.54 per household per day. (See Table 6.9)

Table 6.9

Bidding Game and Site Substitution Results*

	Bidding Game	Site Substitution	
		@$0.04-$0.20	@$0.05-$0.07
Group	Bid ($1976)	($1976)	($1976)
Daytrippers	2.56	1.28-6.39	1.60-2.23
	(2.86)		
Campers	2.48	2.01-10.05	2.51-3.52
	(1.54)		
Population	2.54	1.48-7.40	1.85-2.59
	(2.53)		

Source: Thayer (1981), p. 43, ($1980).

* standard deviations in parenthesis.

Based on this observation, Thayer draws two conclusions. First, that "... the site substitution method, used as a cross check against bidding game (CVM) results, indicates that the survey approach gives reasonable estimates of consumer's welfare loss ..." (Thayer, 1981, p. 43) and, more strongly, that "These results indicate that the (CVM).. can provide accurate (emphasis added) estimates of ... welfare losses associated with environmental degradation". (p. 44) Secondly, and more strongly, Thayer suggests his results "... dispel the argument that inaccurate responses are introduced by the hypothetical nature of the (CVM)." (p. 43)

6. Fisher. Fishers' (1984) paper differs from earlier-reviewed works in that his TCM-CVM comparisons are based on primary research conducted by other researchers. His TCM values are taken from Miller and Hays' (1984) study of consumer surplus values associated with freshwater "fishing days" in five states. CVM values are taken from a study by Loomis (1983) wherein mean estimates of willingness-to-pay (per day) for trout fishing in eleven Western States are estimated. TCM-CVM comparisons can then be made for two states -- Arizona and Idaho -- included in each of the two studies, if we assume that values for "trout fishing" will not differ significantly from values attributable to the more general activity "freshwater fishing".

Relevant values reported by Fisher (1984, pp. 28 and 30) are as

follows:

State	TCM Value for Freshwater Fishing Days (Intra State Mean)	CVM Value for Trout Fishing Days
Arizona	$35.00	$19.54
Idaho	27.00	12.93

Drawing on, and agreeing with, arguments by Brookshire et al. (1982), Fisher argues that CVM values may usefully approximate TCM values notwithstanding "large" differences such as those seen above: "... in comparing the estimates of Loomis with those of Miller and Hays ... the TCM and CVM day values are definitely close enough to eachother that either-- or both -- can serve as a valuable guide to resource managers" (p. 29). Related to the "order of magnitude" issue that will be discussed later in this Chapter, Fisher suggests that ".. if .. information is accurate to within a factor of say, two or three, it (sic) is probably much better in most cases than no information at all." (p. 26)

C. VALUE COMPARISONS: THE CVM AND THE HPM.

The second set of value comparison studies to be considered, focuses on comparisons of values derived by the CVM with those derived from the Hedonic Price Method (hereafter, HPM). The HPM, introduced by Rosen (1974), involves, in operational terms, the identification of 'attributes' associated with a market commodity and the decomposition of the commodity's market price into values attributable to each of the commodity's attributes. In applications of the HPM, the commodity's market price is generally regressed against attributes in efforts to assign values to attributes. Applications of the HPM have been prominent in the literature concerning the value of safety (e.g., Thaler and Rosen, 1975).

There have been three completed studies wherein values for a public good were estimated via the CVM and the HPM. These are the studies by Brookshire, Thayer, Schulze, and d'Arge and by Cummings, Schulze, Gerking, and Brookshire, and by Brookshire, Thayer Tschirhart, and Schulze.

1. Brookshire et al. In the recent study by Brookshire, Thayer, Schulze, and d'Arge (1982) (hereafter BTSd), the public good to be valued via the CVM and HPM was air quality in the Los Angeles metropolitan area. The authors' objective was to use this study "... to validate the survey approach by direct comparison to a hedonic property value study." (p. 165) BTSd develop a theoretical argument for the existence of a rent gradient, which is a mapping onto pollution-commodity space of the differences in housing costs associated with air pollution. They show that the rent differential (dR) can be compared to willingness-to-pay (WTP), and in fact, should serve as an upper bound for WTP values. They also assert that because of the response of the people of California to pollution problems in general, WTP should exceed zero. From this argument, the authors develop and test two hypotheses. The first is that the average WTP for an improvement in air quality over a given community must not be greater than the average rent differential across that community, i.e., $dR \geq WTP$. Second, that average WTP must be strictly positive, i.e., $WTP > 0$.

In order to test these hypotheses, BTSd collected data on air pollution in several communities in Los Angeles. They divided the region into three areas, identifying communities as having poor, fair, or good air quality. A number of independent variables were used to characterize the hedonic rent gradient equation, but they may be characterized by four groups: housing structure variables, neighborhood variables, accessibility variables, and air pollution variables. Due to collinearity between the air pollution measures, two separate log-linear equations were generated, one using nitrogen dioxide (NO_2) as one of the explanatory variables, and the other using total suspended particulates (TSP).

It should be made clear that the rent gradient -- the change (differential) in property values attributable to changes (differential) in air quality -- is the measure to be estimated with the HPM. Thus, BSTd wish to regress housing values against the four groups of variables described above which include air pollution variables; the object, of course, is to identify that part of property value differentials which may be attributed to the site-specific property attribute: air quality. Necessary data for estimating rent gradients were obtained from records concerning 634 home sales during the period January 1977 to March 1978 for nine communities. After estimating the rent gradient, the authors then calculated the rent

differential (dR) for each house in each census tract. The average rent differentials are shown in column 2 of Table 6.10 for the hedonic equation using NO_2 as the pollution variable. The results show monthly rent differentials ranging from $15.44 to $73.78 for air quality improvement from poor to fair, with a sample mean of $45.92. For improvement from fair to good air quality, rent differentials range from $33.17 to $128.46, with a sample mean of $59.09.

For the CVM application, personal interviews of a random sample of 290 households were conducted during March, 1978. In three of the communities, respondents were asked how much they would be willing to pay to improve air quality in their area from poor to fair. In six of the communities, respondents were asked how much they would be willing to pay to improve air quality from fair to good. Respondents were shown maps with isopleths of pollution levels in their area and photographs indicating the visual ranges in poor, fair and good air quality regions. BTSd report that the respondents had little trouble understanding the commodity they were considering. Results of the survey are given in column 4 of Table 6.10. Average monthly willingness to pay (W) for improvement to fair air quality ranges from $11.10 to $22.06, with a sample mean of $14.54. For improvement from fair to good air quality, (W) ranges from $5.55 to $28.18, with a sample mean of $20.31.

Finally, the authors test the two hypothesis noted above. As shown in column 6 of Table 6.10, the calculated t-statistics for the null hypothesis that W = 0, indicate rejection at the 1 percent level in every community. Thus, BTSd conclude that W > 0. In column 7 of Table 6.10, reported t-statistics indicate a failure to reject the null hypothesis that dR > W, at the 10% level. Thus, the a priori hypothesis $0 < W < dR$ developed by BTSd is found to be supported by empirical evidence, a conclusion interpreted by BTSd as "... providing evidence towards the validity of survey methods as a means of determining the value of public goods." (p. 176)

2. Cummings et al. Cummings, Schulze, Gerking and Brookshire (1983) (hereafter CSGB) compare values derived via the CVM with HPM values reported in an earlier paper (Cummings, Schulze, and Mehr, 1978) as they apply to a non-environmental public good: municipal infrastructure in western boomtowns. The authors begin with a discussion of the rationale for using the elasticity measure, (e_1), the elasticity of substitution of wages for municipal infrastructure. The hedonic wage equation used in the Cummings, Schulze, and Mehr (1978) paper is then reviewed. The hedonic elasticity measure (e_1) was based on 209 observations from 26 towns in the Rocky Mountain region. The regression equation resulting from the pooled cross-sectional and time-series was:

$$\ln W = 8.43 + 0.183 \ln D - 0.035 \ln k$$
$$\qquad\qquad\quad (0.022) \qquad (0.017)$$

where W = the wage level

D = the distance from a community to the nearest SMSA

k = the level of a per capita municipal infrastructure

TABLE 6.10 TESTS OF HYPOTHESES FROM THE BTSd STUDY

Community	Property Value Results[a] $\overline{\Delta R}$ (Standard Deviation)	Number of Observations	Survey Results $\overline{W}$ (Standard Deviation)	Number of Observations	Tests of Hypothesis t-Statistics $\mu_{\overline{W}} > 0$[b]	t-Statistics $\mu_{\overline{\Delta R}} > \mu_{\overline{W}}$[c]
(1)	(2)	(3)	(4)	(5)	(6)	(7)
POOR-FAIR						
El Monte	15.44 (2.88)	22	11.10 (13.13)	20	3.78	1.51
Montebello	30.62 (7.26)	49	11.42 (15.15)	19	3.28	7.07
LaCanada	73.78 (48.25)	51	22.06 (33.24)	17	2.74	4.10
Sample Population	45.92 (36.69)	122	14.54 (21.93)	56	4.96	5.54
FAIR-GOOD						
Canoga Park	33.17 (3.88)	22	16.08 (15.46)	34	6.07	5.07
Huntington Beach	47.26 (10.66)	44	24.34 (25.46)	38	5.92	5.47
Irvine	48.22 (8.90)	196	22.37 (19.13)	27	6.08	5.08
Culver City	54.44 (16.09)	64	28.18 (34.17)	30	5.42	11.85
Encino	128.46 (51.95)	45	16.51 (13.38)	37	7.51	12.75
Newport Beach	77.02 (41.25)	22	5.55 (6.83)	20	3.63	7.65
Sample Population	59.09 (34.28)	393	20.31 (23.0)	186	12.02	14.00

[a]Rent differentials for the hedonic housing equation in which $\log(NO_2)$ is the relevant pollution variable are presented here.

[b]The hypotheses to be tested were $H_0: \mu_{\overline{W}} = 0$; $H_1: \mu_{\overline{W}} > 0$. All test statistics indicate rejection of the null hypothesis at the 1 percent significance level.

[c]The hypotheses to be tested were $H_0: \mu_{\overline{\Delta R}} > \mu_{\overline{W}}$, $H_1: \mu_{\overline{\Delta R}} < \mu_{\overline{W}}$. All test statistics indicate that the null hypothesis could not be rejected even at the 10 percent level.

Source: Brookshire, et al., 1982, p. 175.

Standard deviations are shown in parenthesis. Thus, the coefficient on lnk is the measure of the elasticity of substitution of wages for infrastructure:

$$e_1 = -0.035. \quad 3/$$

For the CVM application, a total of 486 residents of Farmington and Grants, New Mexico, and Sheridan, Wyoming 4/, were interviewed. The respondents were first informed of the current level of municipal infrastructure in their area, and the monetary value of the capital facilities. The respondents were then asked how they would reallocate the services provided by their city. Given this reallocation of capital, each respondent was then asked his or her willingness-to-pay for a 10% increase in the city's capital stock, to be allocated in the manner preferred by the respondent. A bidding game was then played until the respondent's maximum WTP was reached. This WTP value, denoted dW, along with an individual's current annual salary (W), was used to calculate:

$$e_{2h} = \frac{\partial W/W}{\%\Delta K}, \quad h = 1, 2, \ldots, 486;$$

where % Δ k is the 10% increase in capital stock. Finally, an average elasticity measure (e_2) was calculated for the individuals in each sample. The results are shown in Table 6.11.

Table 6.11

Elasticity Measures

Hedonic Study*		Survey*					
		Grants		Farmington		Sheridan	
e_1	n	e_2	n	e_2	n	e_2	n
-0.035	209	0.037	115	0.040	278	0.042	93
(0.017)		(0.031)		(0.058)		(0.078)	

Source: Cummings, et al. (1983), pp. 4-6.

* Numbers in parentheses are standard deviations.
n = sample size.

Following a procedure suggested by Scheffe (1970) for comparing a regression coefficient to a sample mean, the authors find the calculated t-statistics to be 0.057, 0.083, and 0.088 for Grants, Farmington, and Sheridan respectively. Against a null hypothesis of equality between e_1 and e_2 (for each of the three towns), the authors report that such low values indicate that one fails to reject the null hypothesis $e_1 = e_2$ at any level of significance. Thus, they conclude that no statistically significant difference between the two measures exits. From this, the authors offer two conclusions. First, their results support the results reported in Brookshire, Schulze, Thayer, and d'Arge (1982) in demonstrating "... that both hedonic and survey approaches yield comparable estimates for the value of selected public goods ...". (Cummings et al., 1983, p. 12) Secondly, the authors suggest that:

> "While interesting, these results do not 'prove' the accuracy of survey measures for public good values; ... survey and hedonic values may be biased vis-a-vis 'true' social values for public goods. There is simply no objective, a priori manner by which the accuracy of survey measures can be 'proven' (or, thus far, disproven ...); if successful, however, repeated experiments of the type reported above may go far in redefining some of the economists' reservations concerning the use of survey methods for valuing public goods." (p. 12)

3. Brookshire et al. (1984). In a recent study by Brookshire, Thayer, Tschirhart and Schulze (hereafter BTTS) an expected utility model of self insurance that incorporates a hedonic price function is presented and applied to low-probability, high-loss earthquake hazards. While the central

focus of the paper is the establishment of a hedonic price gradient for earthquake safety in the Los Angeles and San Francisco areas and a test of the expected utility model, a CVM study was also conducted in Los Angeles which, provides a basis for a comparison of results. The public good of value essentially stems from the Alquist-Priolo Special Studies Zones Act passed by the California legislature in 1972 and amended in 1974, 1975, and 1976. Special Studies Zones are designated areas of relatively elevated earthquake risk as indicated by geologic studies that have identified surface rupture since the Holocene period (approximately 11,000 years ago). Existence of faults, through these geologic studies, may be directly observable through the distortion of physical features such as fences, streets, etc., as well inferred from geomorphic shapes. The total number of SSZ's designated in California as of January 1979 was 251. Of interest is the potential for the Alquist-Priolo Special Studies Zones Act to create a market for avoidance of earthquake risk where no such market existed previous to the passage of the Act. Two elements of the legislation's potential lead to the existence of such a market. First, when an SSZ is designated, property owners are notified thus altering them to an elevation in risk relative to surrounding areas. Second, the process of selling property located in an SSZ requires notification of prospective buyers that in fact the property was located in an area subject to relatively greater earthquake risk.

The impact of the Alquist-Priolo Act through the disclosure requirements form the basis of a testable hypothesis via the HPM. The null hypothesis is that consumers respond to the awareness of hazards associated with SSZ's as illustrated in sales price differentials for homes in and out of an SSZ. The alternative hypothesis being that they do not.

The procedure, data sources and variable structures utilized in estimating the rent gradient for the HPM are those followed in the air pollution study described earlier, (Brookshire et al., 1982). Specific to the earthquake safety attribute a dummy variable which takes on the value 1 for homes in an SSZ and zero otherwise is used in the hedonic equation. Separate equations using housing data for 1972, a period before the Alquist-Priolo Act was passed, and data for 1978, a period after the Act was passed, were estimated. The dummy variable was insignificant in the 1972 equation and significant and of a negative sign in the 1978 equation; indicating that a significant safety variable was in fact a result of the successful enhancement of consumers' awareness of earthquake risk.

In the CVM study, homeowners in and out of SSZ's were asked willingness-to-pay (WTP) and willingness-to-accept (WTA) questions related to the potential transfer of homeownership. Homeowners located in SSZ's were asked how much more they would pay to purchase the same home outside of an SSZ. Homeowners located outside SSZ's were asked how much less expensive their houses would have to be, for them to be willing to relocate in an SSZ.

Utilizing a non-linear specification of the HPM Los Angeles County results indicate that if all other variables in the specification (e.g., housing attributes, etc.) are assigned their mean values, then living outside of an SSZ causes an increase in home value of approximately $4,650 over an identical home located in an SSZ. The CVM results -- the amount that subjects would be willing to pay to purchase the same house outside of an SSZ -- indicates that only 26% of the subjects would be willing to pay some positive amount to move outside of the SSZ. An average of all CVM responses, including zero bids, was $5,920 which is close to the average

sale price differential of $4,650. Homeowners outside an SSZ, when asked how much less expensive their house would have to be to move, responded on average with a value of $28,250.

The results indicate that the WTP measure stemming from the CVM study are quite similar to the HPM. However, the asymmetry between WTA and the WTP is quite striking. The WTP versus WTA dilemma aside, the results suggest a consistent comparison of the HPM and CVM results as applied to earthquake risks.

D. WHAT IS ACCURACY?

Before interpreting the results from comparison studies reviewed above, several comments are in order. Notwithstanding the 'closeness' of comparative values observed by Knetsch and Davis, the above demonstrated notion that CVM-TCM value comparisons generally raise more questions than they resolve, in terms of contributing to assessments of the CVM, should not be surprising. This follows from the myriad of problems with the TCM per se as a method for estimating values for non-market goods. These problems include (Mendelsohn and Brown, 1983; McConnell and Bockstael, 1983, 1984; and Hueth and Strong, 1984): value-allocation assumptions related to multi-purpose 'visits'; dependence of costs on assumptions concerning fixed/variable direct travel costs, costs (benefits?) of time spent in travel and on-site; and problems involved in obtaining values which are appropriately 'marginal' vis-a-vis the site/activity in question. The latter, 'marginal' issue may be best treated by Thayer's site substitution approach (Knetsch and Davis 'willingness to drive' approach). These problems result in the dispelling of what was once regarded as the TCM's greatest potential strength: appealing to the notion that visitor values must equal or exceed travel costs (otherwise, the visit would not be made, see Knetsch and Davis, 1966, pp. 138-140), the TCM must establish a lower bound on 'true' values. While, conceptually , this may be true for simple out-of-pocket travel costs, results from empirical efforts to measure total travel costs seemingly belie this posited 'strength' of the TCM. As demonstrated above, the relationship between TCM values and values derived from the CVM (or any other method) depends, simply, on what is assumed. Thus, Knetsch and Davis find TCM ($70,000) = CVM ($72,000) assuming one-way travel costs valued at 5 cents/mile; the value of time is not addressed. Bishop and Heberlein find TCM ($28.00-plus) > CVM ($21.00) with time valued at one-quarter or more of wage rates. Desvousges, Smith and McGivney (not surprisingly, perhaps, in light of the above), find the TCM value in excess of CVM for deteriorations in water quality and, more remarkably, TCM values less than CVM values for water quality improvements with time valued at full, estimated market rates. Finally, Thayer, abstracting from 'time' issues, finds TCM ($1.28-6.39) < or > CVM ($2.48-2.56), depending on one's estimates for out-of-pocket travel costs.

All else equal, the HPM might be expected to result in value estimates which more closely approximate market values, thereby offering an appealing standard against which CVM values might be compared. Notwithstanding estimation problems in implementing the HPM -- problems which weaken the 'presumption of validity' often accorded methods based on 'real' transactions (Randall et al., 1983, p. 636) -- some bases exist for viewing HPM measures, competently estimated, as minimally providing 'qualitative, order of magnitude', estimates of value. The adjectives 'qualitative, order of magnitude' may describe casual observations as to wage/quality of life trade-offs implied for example, by migrations of workers to Alaska during the construction of the Alaskan pipe-line: some part (hedonic price) of the high reported wages required to attract workers for that project was surely attributable (broadly defined) to environmental amenities. More formally, the results of Ridkers' (1967) seminal work provide compelling empirical evidence of income-environmental trade-offs accepted by individuals: income reductions (hedonic prices) are accepted

(paid) by individuals for quality of life amenities, including environmental amenities.

Estimation problems abound in efforts to implement the HPM -- to name but two: persistent collinearity between 'important' variables and extraordinarily low explanatory power in regression equations (Brookshire et al., 1984). One can only speculate as to the position of estimated HPM values in the range of deviations around a 'true' value for any non-market commodity. In this light, the authors reject as inordinately, and unsupportedly, strong Brookshire et al. 's (1982a) interpretation of results from their comparisons of HPM and CVM values as providing evidence related to the validity (presumably, 'accuracy' vis-a-vis 'true' values) of the CVM as a means for valuing public goods.

One cannot deny, however, the provocativeness of value comparison results reviewed above in section C. Given the differing methodological weaknesses which we understand a priori to be peculiar to each method, the compatability of HPM and CVM measures demonstrated in the four experiments reported in these works is remarkable -- admittedly, it may also be puzzling. Of course, this observation is reminiscent of Randall et al.'s (1983) comment: "Given the relatively weak incentives for careful decision-making in contingent markets ... the relatively strong performance of (the CVM) is perhaps surprising." (Randall et al., 1983, p. 641)

While interesting, surprising, provocative or remarkable, the issue remains as to what one might conclude from the above-reported HPM-CVM value comparisons. Of course, conclusions in this regard require some standard as to accuracy. Thus, our purpose in this section is to reconsider the comparison studies discussed above within a context wherein we first attempt to assess in broad terms the accuracy of each technique. We follow the traditional definition of scientific accuracy which results in statements such as "the measurement is accurate to within ± "x" percent of the measured value". Such a definition of accuracy is essential because estimates of accuracy which economists have implicitly employed, such as the standard error of a regression coefficient in a hedonic equation, do not reflect the many possible sources of inaccuracy such as improper choice of functional form, simultaneous equation bias, or inappropriate assumptions on the distribution of the disturbance term, etc. The only way to incorporate a broader estimate of the total possible range of error is to catalogue the documented range of deviation in measured values for a particular technique. For example, Learner, in an article aptly entitled "Let's Take the Con Out of Econometrics" (Learner, 1983), argues that the only way to assess the true accuracy of econometric estimates is to perform sensitivity analysis over such factors as choice of functional form. Summing up demonstrated possible sources of error as a percent of estimated values then allows determination of an economic equivalent of "reference accuracy".

Reference accuracy is defined as the "limit that errors will not exceed when the device is used under reference operating conditions" (Van Nostrand, 1970, p. 18). Thus, in scientific applications the "device" is a measuring instrument such as a scale used for obtaining weight, whereas in economics the "device" would be an estimation method such as the CVM, TCM, or HPM. "Reference operating conditions" (ROC's), in scientific applications refer to limits on the relevant circumstances under which the measurement is taken such as temperature, atmospheric pressure, etc. In economic applications such as the CVM, limits also exist. For example, in using the CVM, to maintain the hypothetical nature of the survey and avoid strategic bias, the technique possibly should not be employed for current political issues where individuals perceive their answers will influence immediate outcomes (Rowe and Chestnut, 1983).

We will further specify reference operating conditions for the CVM below, but note that, based on discussions given above in Chapters III and IV, the technique must use willingness-to-pay as opposed to willingness-to-accept measures of value and should not be applied to commodities with which people have little or no experience in making prior choices or which involve a high degree of uncertainty.

A second aspect of scientific accuracy, significant digits, should be noted since it is often a point of irritation when non-economists, especially natural scientists, examine benefit estimates produced by economists. An example will make the point clear. An economist might report that the average bid in an application of the CVM was $11.41. the natural scientist will respond that reporting the result in this way is inappropriate since four significant digits are used, which does not reflect the accuracy of the measurement method. In this regard, the standard deviation reported with the average bid is not relevant for assessing accuracy, since a large value can result solely from different individuals having different values (tastes) for the same public good and since a highly biased

average bid may have a small standard deviation. Four alternative ways of reporting the example average bid used above and the implied accuracy of each are as follows:

Number of Significant Digits	Average Bid	Implied Accuracy
4	$11.41	±$ .005
3	$11.4	±$ .05
2	$11	±$ .50
1	$1 x 10^1	±$5.00

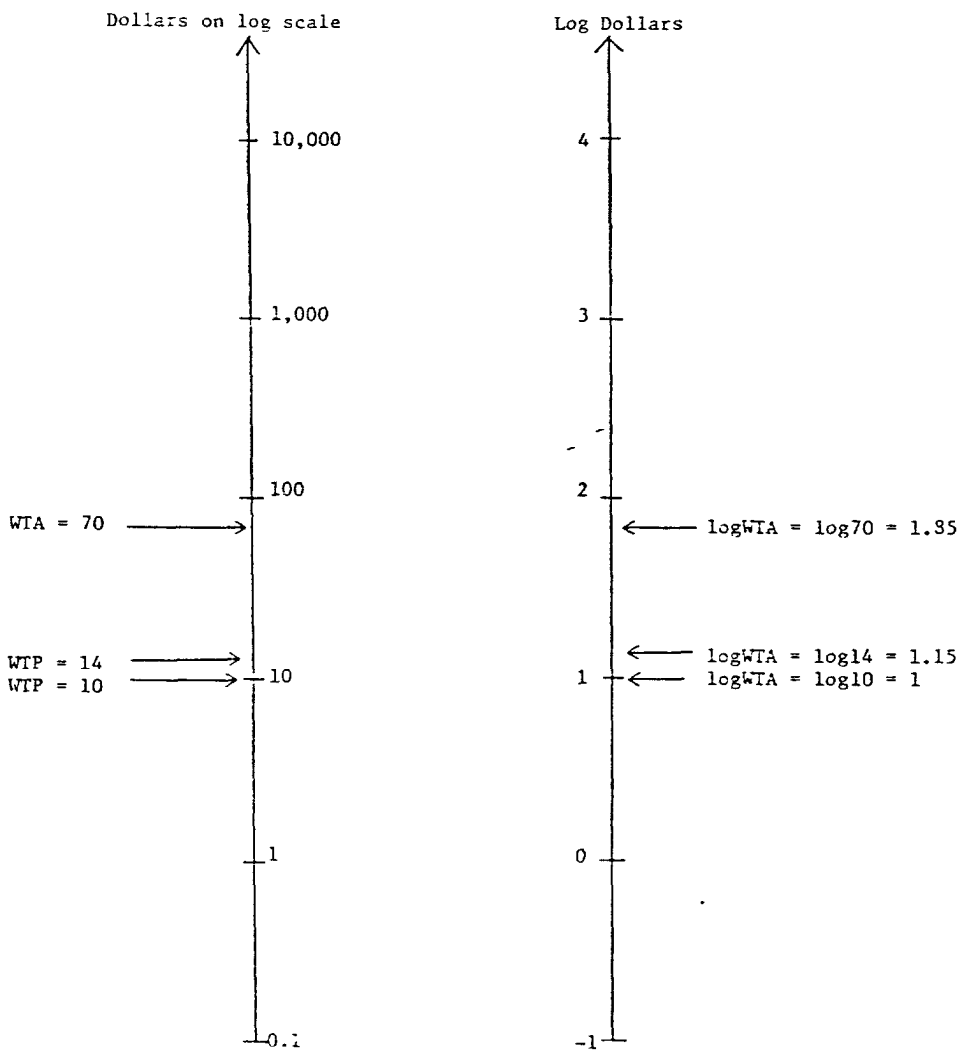
Note that the implied accuracy is one half of the value of the last reported digit. (Kreyszig, 1979, p. 758) Economic value estimates are almost always reported as though they have at least three significant digits. We will argue below that they, in fact, have a level of accuracy which implies no more than one significant digit, i.e., an accuracy no better than about ±50 percent of the measured value.

A third view of the accuracy of scientific measurements relates to the "order of magnitude" of the estimate. For example, a scientist may argue that the amount of CO_2 gas dissolved in the earth's oceans (an important quantity in estimating the likelihood that burning fossil fuels will alter the earth's climate through the greenhouse effect) is only known to within one order of magnitude. What this would imply for estimating the accuracy of economic measures is shown on the vertical scale in Figure 6.1, which is logarithmic in that each unit of distance on the scale, moving from bottom to top, implies a tenfold increase in magnitude. Based on discussions given above, a willingness-to-pay bid of $10 obtained using the CVM payment card might be raised by 40% to $14 by applying iterative bidding. A corresponding willingness-to-accept bid may be as much as five times greater than the WTP measure, or $70.00. The arrows in Figure 6.1 illustrate these example bids along the logarithmic scale. Note how the $10 and $14 bids are relatively close, "of the same order of magnitude", while the $70 bid is close to the $100 level on this scale, an order of magnitude larger than the previous two bids. Thus, one might argue that the iterative and non-iterative willingness-to-pay bids are "close", of the same order of magnitude, while hypothetical willingness-to-pay and hypothetical willingness-to-accept measures are not "close" and may differ by about one order of magnitude. Physical scientists and health scientists often argue that "order of magnitude" estimates are the best that can be made for complex environmental processes which may be relevant for many benefit-cost studies. As a result, economists may be in a relatively comfortable position if they can avoid errors as large as one order of magnitude such as implied by the difference between hypothetical willingness-to-pay and willingness-to-accept measures of value.

The range of possible error for the CVM derived from selected sources of bias is seen in Rowe et al. (1980). Rowe et al. state that in examining the effects of starting point, vehicle, information, and strategic bias, as reported in several studies reviewed by them, only strategic bias did not seem to have a significant affect on bids. They conclude that the sum of starting point, vehicle and information bias can be as large as 40 percent of the estimated value. One additional source of bias is relevant. Schulze

Figure 6.1

Order of Magnitude Estimates



et al. (1981) show that use of a payment card to record bids, results in bids as much as 40 percent lower than obtained with the use of iterative bidding. Even though, based on the experimental evidence of Chapter IV, we reject outright hypothetical willingness-to-accept measures of value, the sum of the demonstrated possible biases is about 64 percent. In other words, an upper bound bid of $10 could be reduced to $6.00 by the sum of the effects of starting point, vehicle and information bias and further reduced to $3.60 by choice of a payment card for collecting bids. Averaging $10.00 and $3.60 gives an example midpoint bid of $6.80. If we report this bid, $6.80 as having an accuracy of ±50 percent the implied range would be $10.20 to $3.40, very close to the range implied by known potential biases in the CVM. Thus, one might tentatively conclude that, given the current state of the arts, the CVM is not likely to be more accurate than ±50 percent of the measured value.

How accurate are the HPM and the TCM? Unfortunately, detailed estimates of the possible sources for and magnitudes of errors associated with these techniques, are not available. Even though HPM and TCM (indirect market) techniques are regarded by some as yielding accurate, market-analogous values, a large number of theoretical and econometric issues are relevant to their use in estimating values for public goods. For example, a possible identification problem which may arise in the use of indirect market methods for value estimation has been analyzed by Brown and Rosen (1982). As noted above, a special problem exists with respect to assumptions made concerning the value of time spent in travel when willingness-to-pay estimates are derived using the TCM (see for example, Cesario, 1976; Mendelsohn and Brown, 1983). All of these problems suggest that estimating willingness-to-pay values for environmental commodities via indirect market methods may well involve sources for errors that exceed, in substance and number, those relevant for estimates of ordinary demand equations for market goods. However, we can show that even estimation of ordinary demand equations is subject to surprisingly large errors. Since no systematic study has been done of the possible errors in indirect market methods, we will assume that the errors in these methods are at least as large as those which can be shown to exist for estimates of market demand.

Coursey and Nyquist (1983) apply a number of estimation techniques which allow for alternative assumptions about residual distributions Of errors (including least squares, least absolute errors, Huber, Cauchy, exponential power and student's t) in estimating demand equations for six market commodities in three different countries. Thus, 18 separate demand equations were estimated using six different procedures. Strong evidence was found that the assumption of normality on the disturbance term was generally violated and that the use of robust alternatives to "normality" assumptions was appropriate. Further, estimates of the intercept, income elasticity and own-price elasticities in each case were highly sensitive to choice of estimation technique. Changes in estimated intercepts from the use of different techniques varied from 5 to 747 percent and exceeded 50 percent in 8 of the 18 demand equations. Changes in estimated income elasticities across techniques varied from 3 to 851 percent and exceeded 50 percent in 5 of the 18 demand equations. Finally, changes in estimated price elasticities ranged from 14 to 183 percent across techniques with a change greater than 50 percent in 12 of the 18 demand equations.

A few calculations will show that even if initial price and quantity are equal, variations in estimated price elasticity like those commonly

found in the Coursey and Nyquist study will result in variations in estimated willingness-to-pay which are greater than ±50 percent. For example, for the United States, the estimated price elasticity of demand for clothing varies from about -.05 to -1. For a 20 percent increase in quantity, the ratio of upper to lower bound estimates of willingness-to-pay is then about 3.2 assuming that the price elasticities are constant. A 3 to 1 ratio is, of course, consistent with an error range of ±50 percent. It would be most useful if we had information as to the sensitivity of measures estimated by indirect market methods to the use of alternative functional forms and alternative included variables as well as the relevance of simultaneous equation bias and alternative assumptions on the disturbance term. However the potentially large errors in estimating the parameters of ordinary demand equations, discussed above, would seem to suggest that the accuracy of values estimated with indirect market methods is likely to be no better than ±50 percent.

If errors in the CVM and the two indirect market methods, HPM and TCM, are likely to limit accuracy to no better than ±50 percent of measured values, what are the implications of the comparison studies? If, for example, the measured value for a particular commodity using the CVM IS $10.00 and the same commodity, under the same circumstances is valued at $28.00 using the TCM, are the two measures different? Many of the authors of the comparison studies would argue that these measures are not only different but, that since the TCM is based on actual as opposed to hypothetical behavior, it must be the correct value. In contrast, one might argue that, based on the analysis of accuracy presented above, these two example values are not distinguishably different since the CVM value has a range of at least $5 - $15 and the TCM value has a range of at least $14 - $42 and these two ranges overlap.

Table 6.12 presents a summary of results from the comparison studies reviewed earlier in this chapter. Some of these studies offer a range of values for the valuation methods employed based on calculated variances, standard errors, etc.; however, in none of the studies does one find considerations relevant for the "reference accuracy" of measures associated with their estimation techniques. The most striking aspect of data in Table 6.12 is that of the 75 comparisons given for the 7 studies, none of the comparison studies show a significant difference between values drawn from alternative techniques using our criterion for accuracy. In other words, if reference accuracy is expressed in terms of ±50 percent, ranges for reference accuracy for the CVM and indirect market methods overlap in 13 of the 15 cases (excluded are Desvousges, et al. (a) and Brookshire et al. (1982) (b). This finding of a lack of a significant difference between CVM and indirect market values extends to Brookshire et al.'s case (a) if reference accuracy is stated in terms of ±52%, and to Desvousges et al.'s case (a) when reference accuracy is expressed in terms of ±60%. Thus, in the 50-60% range -- surely a palatable range given the ±50% range of error attributed to estimates of ordinary demand relationships -- CVM values are consistently "accurate" estimations for values derived with indirect market methods.

The reader may easily draw an incorrect conclusion at this point. This result does not establish the accuracy of CVM measures for any particular commodity. Rather, it simply appears that values derived from the CVM fall within the range of "reference accuracy" (given the admittedly large error bounds developed above) for those commodities where indirect market measures

TABLE 6.12

SUMMARY OF RESULTS FROM COMPARISON STUDIES

Study	CVM RESULTS	Value 1/	INDIRECT MARKET STUDY Method	Value 1/
Knetsch and Davis (1966)	Recreation Days	$1.71 per household/day	TCM	S1.66 per household/day
Bishop and Heberlein (1979)	Hunting Permits	$21.00 per permit	TCM value of time=0 value of time=¼ median income value of time=½ median income	 $11.00 $28.00 $45.00
Desvousges, Smith and McGivney (1983)	Water Quality Improvements: a) loss of use b) boatable to fishable c) boatable to swimmable	User Values: 2/ average (across question format) $21.41 $12.26 $29.64	TCM	user values: $82.65 $ 7.01 $14.71
Seller, Stoll and Chavas (1984)	Boat Permit to: Lake Conroe Lake Livingston Lake Houston	close-ended Consumer Surplus: $39.38 $35.21 $13.01	TCM	Consumer Surplus: $32.06 $102.09 $13.81
Thayer (1981)	Recreation Site	Population Value per household per day: $2.54	Site Substitution	Population Value per household per day: $2.04
Brookshire, et al. (1982)	Air Quality Improvements: a) poor to fair b) fair to good	monthly value 3/ $14.54 $20.31	HPM (property values)	monthly value: $45.92 $59.09
Cummings, et al. (1983)	Municipal infrastructure in: a) Grants, N.M. b) Farmington, N.M. c) Sheridan, WY	elasticity of substitution of wages for infrastructure -0.037 -0.040 -0.042	HPM (wages)	elasticity of substitition of wages for infrastructure; 29 municipalities: -0.035
Brookshire, et al. (1984)	Natural Hazards (earthquakes) information	$47 per month	HPM (property values	$37 per month

1/ Mean values amongst respondents.
2/ Values apply to post-iteration bids for users of the recreation sites.
3/ Values for sample population.

can be obtained. Assuming that, within the range of ±50%, value estimates derived from indirect market methods include "true" valuations by individuals, these results suggest that CVM values may yield "accurate" estimates of value in cases where individuals have had some opportunity to make actual previous choices over that commodity in a market framework. These studies do not demonstrate that people are capable of providing market like values using the CVM for commodities which are not already being traded in existing markets, at least to a limited or indirect degree. In this latter regard, examples include such "commodities" as existence and option values for preserving an environmental asset over which people have no experience in making prior choices. We will examine this argument in greater detail below.

E. IMPLICATIONS FOR ASSESSMENTS OF THE CVM.

If, as suggested above, the CVM is indeed "accurate" vis-a-vis estimates for individual values derived from indirect market methods, we must then inquire as to the general implications of this observation for one's assessment of the CVM. In this regard, we are left with the necessity of defining conditions -- "reference operating conditions" (ROC) -- relevant for estimation methods which may be expected to yield value measures which satisfy the criterion of reference accuracy. To this end, we begin by considering ROC's implied by the institution underlying indirect market methods: the market.

In our society "the market" consists of many amorphous "markets" which differ in such things as degrees of organization and the necessity for negotiation. Thus, as observed by Knight (1951):

> "In economics (a market) means the whole area, often indefinitely defined, within which buyers and sellers of a commodity come together and fix a common price .... The wheat market is practically the world ... the market for ... brick from a small factory may not extend beyond a few miles." (p. 68)

As further examples in these regards, the market for groceries is relatively well organized and exchange involves little if any negotiation. Towards another end of the spectrum, the market for used furniture is less well organized and exchange can, in some settings (e.g., the flea market), involve considerable negotiation.

Also of importance for our consideration is the fact that economic deductions drawn from "the market" are complicated by the fact that commodities traded in a market are often heterogeneous. Thus, Knight asks: "... is wheat in Paris the same commodity as wheat in Chicago? ... is a physically equivalent ... can of peas with a label which is a guarantee of quality, effectively the same commodity as if it had an unknown name?" (p. 69) In terms of the efficacy of the market vis-a-vis fixing "a common price", these complexities are substantively increased when dissimilar commodities are jointly offered. An example might be a house; to paraphrase Knight, are two physically equivalent (floor space, rooms, paint, appliances, etc.) houses, one located in (e.g.) neighborhood A and one in neighborhood B, the same commodities? Most often, the answer is "no" inasmuch as other neighbor-related "commodities" are offered in joint supply with the house: crime rates, quality of schools, proximity to beaches, theaters, etc., and, possibly, environmental (air) quality. Each of these commodities, in most cases valued and desirable in their own right, are obtained only in the housing "package". Since one cannot, in choosing a house, pick the crime rate from one neighborhood, the school system of another and air quality from still another, the implicit market valuation of these commodities -- "attributes" of the house in a given neighborhood -- will be imperfect measures of "true" values associated with these attributes.

Whatever the characteristic of any given market, one of the most important characteristics of the set of interrelations involving the process of competing bids and offers which we call "the market" is its capacity to "... generate high quality information at low cost." (Heyne,

1983, p. 125) Thus, "... the most important single cause of exceptions to (market laws) ... is found in the condition: people do not know the facts". (Knight, 1951, p. 69) The better organized the market, the better that people will "know the facts". In these regards, prices provide valuable information and "... the more such prices there are, the more clearly and precisely they are stated and the more widely they are known, the greater will be the range of opportunities available to people". (Heyne, 1983, p. 125).

Thus, key "reference operating conditions" (ROC's) relevant for the market institution include; first, the process of competing bids and offers which generates experience -- familiarity -- with that process; secondly, and implied by the preceeding, the generation of information via repeated trials whereby again, experience and familiarity with commodities and exchange are derived; and thirdly, incentives for an individual's acquiring and "processing" information imposed by his/her limited income juxtaposed with a more or less strong desire to maximize consumption/savings opportunities (maximizing behavior).

The importance of the ROC's described above is made manifest in experimental economics wherein efforts are made to simulate these conditions in an experimental setting. In Smith's (1982) recent experiments with auction mechanisms for public goods the following rules (institution) are imposed: (1) subjects offer bids within a well-defined information context which allows subjects to calculate their net (monetary) gains; (2) repetitive trials are required, which, along with a veto mechanism, provide experience and familiarity -- the opportunity to learn maximizing strategies; (3) rules for group equilibrium are defined (in this case, unanimous agreement). (Smith, 1984, p. 927) Aside from Smith's work, results from experimental economics in general-make clear the importance Of market-like incentive structures and the trial-feedback-learning process in any effort to form incentive compatible institutions and/or, more importantly, to elicit true, market-like preference revelation. As noted in Smith's work, the importance of repetitive trials -- a sequence of trials whereby the individual 'learns' optimal strategies appropriate for the new institution -- is further reflected in Coppinger et al.'s (1980) observation: "(one may) question whether any meaningful one-shot observation can (therefore) be made on processes characterized by a dominant strategy equilibrium". Moreover, we know from our discussions in Chapters IV and V that efforts to simulate the market institution require that elicitation modes focus on WTP (as opposed to WTA) measures and that there be little uncertainty associated with outcomes of bidding processes.

From the above, we may deduce the following ROC's relevant for the CVM.

1) subjects must understand, be familiar with, the commodity to be valued.
2) subjects must have had (or be allowed to obtain) prior valuation and choice experience with respect to consumption levels of the commodity.
3) there must be little uncertainty,
4) WTP, not WTA, measures are elicited.

ROC's 1 and 2 derive directly from the market institution (which provides high quality information at low cost). Moreover, in terms of ROC 1,

results from psychological research (Chapter V, above) point to distortions in decision processes (framing biases, etc.) that arise when individuals are unfamiliar with decision contexts; regarding ROC 2, results from experimental economics emphasize the importance of iterative trials which serve to provide subjects with valuation and choice experience -- subjects must "learn" maximizing strategies; ROC 3 derives directly from research in psychology and experimental economics: under conditions of uncertainty, valuation decisions may be subject to distortions resulting from the use of a wide range of heuristic devices. Finally, as discussed above in Chapters III and IV, WTA measures are generally found to be highly distorted vis-a-vis "true" valuations as a possible result, psychologists would argue, of cognitive dissonance.

The relevance of the above-described ROC's lies in our expectation that, if the CVM institution satisfies them, we would expect the resulting measure of value to approximate market-analogous values within a range of error defined by "background" sources of error, suggested at the present time to be no less than ±50 percent. If ROC's are not satisfied, the range of reference accuracy increases, reflecting the errors associated with the excluded ROC.

A major state-of-the-arts problem is that we know little about the errors associated with the Reference Operating Conditions (Table 6.13). Received research results suggest that if WTA measures are used rather than WTP measures, the WTA measure may be 3 or more times larger than WTP. In terms of ROC's 1-3, however, we lack the data that would allow us to quantify reference accuracy. As noted above, results from psychological and experimental economics research tell us only in qualitative terms that distortions -- errors -- will result when these ROC's are unsatisfied.

TABLE 6.13

REFERENCE OPERATING CONDITIONS AND IMPLIED REFERENCE ACCURACY

	Reference Operating Condition	Implied Reference Accuracy If ROC Not Satisfied
1.	Familiarity With Commodity	unknown
2.	Valuation/Choice Experience	unknown
3.	Little Uncertainty	unknown
4.	WTP Measure	at least 300%

In Table 6.14, data are given concerning the extent to which ROC's were generally satisfied in selected applications of the CVM; these applications are described in considerable detail above and in Chapter III. Thus, in Brookshire et al.'s study of air quality in Los Angeles, subjects were clearly familiar with the commodity, "smog"; with average turn-over of housing in the L.A. area of 3 years, subjects generally can be assumed to be knowledgeable of the air quality attribute related to housing and housing costs (advertisements for housing in the L.A. newspaper will many times include a description of air quality), in which case subjects had some degree Of experience in valuing choices with respect to "consumption levels" of the commodity (improved air quality). Also, uncertainty played a negligible role in Brookshire et al.'s CVM application wherein WTP measures were elicited. Analogous arguments apply to the study of municipal infrastructure by Cummings et al.

To generalize these observations, we can identify eight studies which, to differing degrees, essentially satisfy the above described ROC's: those given in Table 6.12. In each of these studies, indirect market measures of value (using either the TCM or the HPM) were derived in addition to value measures derived by the CVM. As indicated above, using ±50 percent for reference accuracy, in each of the eight cases we would fail to reject the hypothesis that the CVM measures and the indirect market measures are the same. If one accepts Hedonic (or Travel Cost) measures as including, within a ±50 percent range for reference accuracy, values which reflect market-analogous revelations of preferences, then one's acceptance of the accuracy of CVM values for applications wherein the ROC's are satisfied turns on the question: do the fifteen comparisons given in these eight studies constitute the preponderance of evidence required in science to establish "facts"?

Finally, we must ask: what of the CVM studies which do not satisfy one or more of the ROC's -- particularly ROC's 1-3 about which we know little in terms of reference accuracy (e.g., referring to Table 6.14 the study designed to derive existence and option values for visibility in the Grand Canyon by Schulze et al. and Burness et al.'s toxic waste

TABLE 6.14

ROC'S SATISFIED IN SELECTED APPLICATIONS OF THE CVM

REFERENCE OPERATING CONDITION	CVM EXPERIMENT: Brookshire, et al. (air quality, Los Angeles	Cummings, et al. (Municipal Infrastructure in boomtowns	Schulze, et al. (Visibility in Grand Canyon	Burness, et al. (toxic wastes)
1) Familiarity With Commodity	YES	YES	NO	NO
2) Valuation/Choice Experience	YES	YES	NO	NO
3) Little Uncertainty	YES	YES	YES(?)	NO
4) WTP Measure	YES	YES	YES	YES

study). In such cases we can say no more than that there exists no positive evidence that would support the accuracy of such measures vis-a-vis market or market-related values. It must be said, however, that negative evidence in this regard does exist. Order of magnitude differences between initial valuations and valuations derived after prior experience (from iterative trials) with choice mechanisms are suggested by research in experimental economics. Research in psychology has firmly established the distortions in choices which attend decision environments characterized by uncertainty and unfamiliar learning/decision contexts. In short, we can neither confirm nor deny the accuracy of CVM values derived in applications which do not satisfy the ROC's; given the present state of the arts. However, available evidence suggests that such measures may be seriously distorted.

F. FINAL REMARKS

The seven chapters of Part I of this book have focused on three major issues relevant for the CVM. First, an effort was made to provide the reader with some flavor for how and why interest in the CVM was initiated as well as the rationale for and nature of early experimental efforts to develop the method; these were the topics addressed in Chapters I and II. Secondly, the authors surveyed the literature to the end of identifying claims for sources of bias in value measures derived with the CVM, after which the authors drew on research results reported in the economics and psychology literature in efforts to assess the potential nature and importance of these biases; our efforts to assess the strengths and weaknesses claimed for the CVM were the substance of Chapters III through V. Third, and finally, in this Chapter -- Chapter VI -- the authors have attempted to focus the results of earlier analyses on the question of central interest in this book: how might one assess the accuracy of measures derived with the CVM, and what are the implications of such an inquiry for the state of the arts of the CVM as a means for valuing non-market, public goods?

Before summarizing results from the authors' considerations of this state of the arts question, the reader is reminded of the ultimate end sought in this work, viz, a broad, profession-wide evaluation of the CVM. Something akin to this broad assessment of the CVM is sought in the Conference described in Chapter I at which the state of the arts question is to be considered by several scholars involved in one way or another with the CVM as well as by a Review Panel consisting of outstanding scholars in the economics and psychology professions. Thus, the authors offer no "conclusions" per se at this time. We have suggested a framework for assessing the accuracy of CVM measures which will hopefully be found as provocative in the Conference's collective considerations of the CVM. The following summary of the authors' arguments are offered within this context. The response to this assessment framework by Conference participants will be described in Part II, and efforts to draw final conclusions as to the state of the arts of the CVM will be given in Chapter XIV.

Our approach to assessing the state of the arts of the CVM is couched in terms of instruments and scientific measuring systems wherein "accuracy" is defined as follows: "... conformity of an indicated value to an accepted standard value, or true value ... accuracy should be assumed to mean reference accuracy..." (Van Nostrand, 1970, p. 17). Reference accuracy, expressed in terms of a range or span around the measured variable (measure ±X%), defines the limits that errors will not exceed when a measure is obtained under Reference Operating Condition. Since our accepted standard, or true values, are market values, the ROC's for the CVM suggested by the authors are drawn from what we know of the market institution, as well as what has been learned in analyzing market-like behavior in experimental economics and in psychology-related research. These suggested ROC's are:

1) subjects -- participants in the CVM -- must understand, (be familiar with) the commodity to be valued.
2) subjects must have had (or be allowed to obtain) prior valuation and choice experience with respect to consumption levels of the commodity.
3) there must be little uncertainty.

4) WTP, not WTA, measures are elicited.

Ideally, experimental research would have defined limits on errors associated with applications of the CVM which fail to satisfy any one of the ROC's. This is not the case, however. In the present state of the arts, such limits (very large limits) are known only in terms of ROC 4: WTA measures may approximate market values only in a range of some ±300% -- plus!

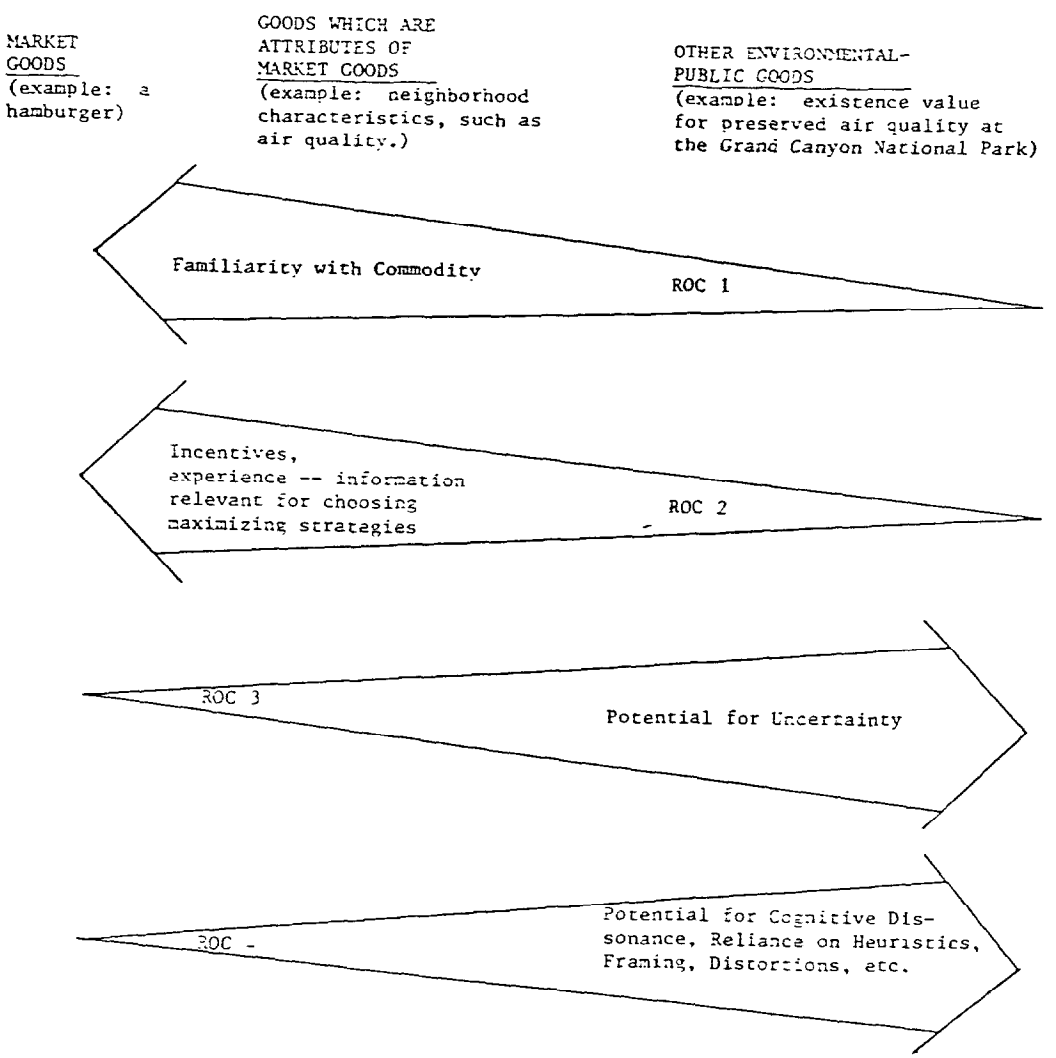
In considering indirect market values -- values estimated by the TCM and HPM -- we assert that reference accuracy for these measures can be expected to be no better than that for estimates of parameters of ordinary demand functions (which arise from assumptions on residual distributions), which is the measured value plus-or-minus 50%. State of the arts information allows one to go beyond simply deducing ROC's for the CVM and, essentially, asserting that CVM applications which satisfy the ROC's will yield reference-accurate measures. Eight studies have been identified (Table 6.12 above) which derive CVM values as well as values from indirect market methods and which satisfy the ROC's for the CVM. In each case, one fails to reject the hypothesis that the CVM measure is the same (in reference accuracy terms) as the indirect market measure. Thus, if one accepts the reference accuracy of ±50% as including "true" market values, one has six tests which consistently infer that Reference Accuracy measures derived from the CVM are "valid". Whether or not these six cases constitute the preponderance of evidence required in the scientific method to establish "facts" is, of course, a matter of judgement.

One may find little comfort in these observations in terms of the general promise of the CVM as a means for estimating "accurate" values attributable to broad categories of public/environmental goods. This follows from the fact that, given the present state of the arts, a limited number of environmental "commodities" are amenable to CVM applications, where the ROC's are satisfied. For such applictions, where the ROC's are not satisfied, the present state of the arts does not allow us to conclude that accurate or inaccurate measures will result. It must be said, however, that while positive evidence vis-a-vis the accuracy of CVM measures derived under these circumstances does not exist, considerable negative inferential evidence does exist in this regard.

In closing, the authors recognize that while an assessment framework based on reference accuracy and the Reference Operating Conditions may in form parallel objective frameworks for assessing accuracy in other sciences, it may fall well short of "objectivity" vis-a-vis assessments of the CVM. This follows from the obvious fact that while the ROC's per se may be objectively deduced from market institutions, their application to assessments of a CVM study may generally be subjective. For example, one may ask: what degree of "familiarity" with a commodity is required to satisfy ROC 1; how much value/choice experience (or how many repetitive trials) is (are) required to satisfy ROC 2; and how much is "little uncertainty" (ROC 3)? In response to these questions, our knowledge of markets, lessons drawn from experimental economics and psychological research tell us little more than that, in moving from pure public goods to common market goods, we can expect something of a continuum in meeting ROC's as exemplified in Figure 6.2. Thus, moving from an "existence value" to a hamburger, we expect individuals to be increasingly familiar with the "commodity" and to have had greater market-related experiences; along this

Figure 6.2

ROC's and Market, Non-Market Commodities



continuum, uncertainties as to outcomes of transactions and the potential for problems related to cognition are reduced.

In efforts to deal with these issues, the state of the arts is one wherein we can simply say that evidence exists which supports the proposition that indirect market experience with a commodity may serve to satisfy the ROC's: when the environmental good is a distinct attribute Of a market-related good (water quality in a time/travel cost recreation trip or air quality as an attribute of housing locations/costs), experience/familiarity with the market good seemingly spills over to the individual's ability to value the attribute. Thus, while not totally answering the "what degree" and "how much" questions regarding the satisfaction of ROC's, comparison studies may suggest classes of environmental/public goods which may be taken a priori as those which would satisfy the ROC's for the Contingent Valuation Method.

ENDNOTES

Chapter VI

1) While Rosen may be credited with the initial, rigorous theoretical development of the HPM, the HPM *per se* was used in earlier studies, most prominently in Ridker, 1967.

2) Researchers at the University of Wyoming have developed data amenable to CVM and HPM analysis related to ozone concentrations in Southern California; drafts of final comparative results are unavailable at the time of this writing, however.

3) Although the authors do not discuss the robustness of these results, performing simple two-tailed tests on the coefficient on lnk -- where the null hypothesis is that it is not significantly different from zero -- the null hypothesis is rejected at the 5% level. Thus e_1 is negative and significantly different from zero.

4) These towns were included in 26 towns from which data were used in the HPM study.